09-05-00 A

Box SEQ

09/01/00 JC924 U.S. PTO

Certificate of Mailing

Date of Deposit September 1, 2000 Label Number: EL509049835US

I hereby certify under 37 CFR 1.10 that this correspondence is being deposited with the United States Postal Service as **"Express Mail Post Office to Addressee"** with sufficient postage on the date indicated above and is addressed to: BOX PATENT APPLICATION, Director for Patents, Washington, D.C. 20231.

Luis A. Cruz
Printed name of person mailing correspondence

Signature of person mailing correspondence

09/01/00 JC675 U.S. PTO 09/654743

UTILITY PATENT APPLICATION TRANSMITTAL UNDER 37 CFR §1.53(b)	
Attorney Docket Number	07891/003005
Applicant	ROBERT G. KORNELUK, ALEXANDER E. MACKENZIE, STEPHEN BAIRD, AND PETER LISTON
Title	MAMMALIAN IAP GENE FAMILY, PRIMERS, PROBES, and DETECTION METHODS

PRIORITY INFORMATION:

This application is a continuation of and claims priority from United States patent application 08/576,956, filed December 22, 1995; which is a Continuation-in-Part of United States patent application 08/511,485, filed August 4, 1995, now issued as U.S. Patent No. 5,919,912.

APPLICATION ELEMENTS:

Cover sheet	1 page
Specification	88 pages
Claims	6 pages
Abstract	1 page
Drawing	50 sheets
Combined Declaration and POA, which is: □ Unsigned; □ Newly signed for this application; ■ A copy from prior application 08/576,956 and the entire disclosure of the prior application is considered as being part of the disclosure of this new application and is hereby incorporated by reference therein.	2 pages
Sequence Statement	2 pages
Sequence Listing on Paper	42 pages
Sequence Listing on Diskette	1 diskette
Small Entity Statement, which is: □ Unsigned; □ Newly signed for this application; ■ A copy from prior application 08/576,956 and such small entity status is still proper and desired.	2 pages

Preliminary Amendment	16 pages
IDS	2 pages
Form PTO 1449	5 pages
Cited References	0 references
Recordation Form Cover Sheet and Assignment	0 page
Assignee's Statement	0 page
English Translation	0 page
Certified Copy of Priority Document	0 page
Return Receipt Postcard	1

FILING FEES:	
Basic Filing Fee: $345	$345.00
Excess Claims Fee: 47 - 20 = 27 x $9	$243.00
Excess Independent Claims Fee: 16 - 3 = 13 x $39	$507.00
Multiple Dependent Claims Fee: $260/$130	
Total Fees:	$1095.00

■ Enclosed is a check for $1095.00 to cover the total fees.
□ Charge [**AMOUNT**] to Deposit Account No. 03-2095 to cover the total fees.
□ The filing fee is not being paid at this time.
■ Please apply any other charges, or any credits, to Deposit Account No. 03-2095.

CORRESPONDENCE ADDRESS:

Kristina Bieker-Brady, Ph.D.
Reg. No. 39,109
Clark & Elbing LLP
176 Federal Street
Boston, MA 02110
Telephone: 617-428-0200
Facsimile: 617-428-7045

Signature James DeCamp Reg. No. 43,580

9/1/00
Date

ATTORNEY DOCKET NO. 07891/003001

Applicant or Patentee: Robert G. Korneluk et al.
Serial or Patent No.: 08/576,956
Filed or Issued: December 22, 1995
For: MAMMALIAN IAP GENE FAMILY, PRIMERS, PROBES, AND DETECTION METHODS

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS
(37 CFR 1.9(f) and 1.27(d)) - NONPROFIT ORGANIZATION

I hereby declare that I am an official empowered to act on behalf of the nonprofit organization identified below:

Name of Organization: University of Ottawa
Address of Organization: 550 Cumberland, Ottawa, Ontario, Canada K1N 6N5
Type of Organization:

[X] UNIVERSITY OR OTHER INSTITUTION OF HIGHER EDUCATION
[ ] TAX EXEMPT UNDER INTERNAL REVENUE SERVICE CODE (26 USC 501(a) and 501(c)(3))
[ ] NONPROFIT SCIENTIFIC OR EDUCATIONAL UNDER STATUTE OF STATE OF THE UNITED STATES OF AMERICA
(NAME OF STATE:)
(CITATION OF STATUTE:)
[ ] WOULD QUALIFY AS TAX EXEMPT UNDER INTERNAL REVENUE SERVICE CODE (26 USC 501(a) and 501(c)(3)) IF LOCATED IN THE UNITED STATES OF AMERICA
[ ] WOULD QUALIFY AS NONPROFIT SCIENTIFIC OR EDUCATIONAL UNDER STATUTE OF STATE OF THE UNITED STATES OF AMERICA IF LOCATED IN THE UNITED STATES OF AMERICA
(NAME OF STATE:)
(CITATION OF STATUTE:)

I hereby declare that the nonprofit organization identified above qualifies as a nonprofit organization as defined in 37 CFR 1.9(e) for purposes of paying reduced fees under section 41(a) and (b) of Title 35, United States Code with regard to the invention entitled MAMMALIAN IAP GENE FAMILY, PRIMERS, PROBES, AND DETECTION METHODS by inventor(s) Robert G. Korneluk, Alexander R. MacKenzie, and Stephen Baird described in

[ ] the specification filed herewith.
[X] application serial no. 08/567,959, filed December 22, 1995.
[ ] patent no. , issued .

I hereby declare that rights under contract or law have been conveyed to and remain with the nonprofit organization with regard to the above identified invention.

If the rights held by the nonprofit organization are not exclusive, each individual, concern or organization having rights to the invention is listed below* and no rights to the invention are held by any person, other than the inventor, who could not qualify as a small business concern under 37 CFR 1.9(c) or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d) or a nonprofit organization under 37 CFR 1.9(e).

*NOTE: Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities. (37 CFR 1.27)

Full Name: Apoptogen, Inc.

Address: 100 International Blvd., Etobicoke, Ontario, Canada M9W 6J6

[ ] INDIVIDUAL [X] SMALL BUSINESS CONCERN [ ] NONPROFIT ORGANIZATION

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

Name: Jean Farrall

Title: Director, Research Services

Address: University of Ottawa, 115 Seraphin Marion, Ottawa, Canada

Signature: Jean Farrall **Date:** 15 March 1996

ATTORNEY DOCKET NO. 07891/003001

Applicant or Patentee: Robert G. Korneluk et al.
Serial or Patent No.: 08/576,956
Filed or Issued: December 22, 1995
For: MAMMALIAN IAP GENE FAMILY, PRIMERS, PROBES, AND DETECTIONS METHODS

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS
(37 CFR 1.9(f) and 1.27(c)) - SMALL BUSINESS CONCERN

I hereby declare that I am

[X] the owner of the small business concern identified below:
[ ] an official of the small business concern empowered to act on behalf of the concern identified below:

Name of Small Business Concern: Apoptogen, Inc.

Address of Small Business Concern: 100 International Blvd., Etobicoke, Ontario, Canada M9W 6J6

I hereby declare that the above identified small business concern qualifies as a small business concern as defined in 13 CFR 121.12, and reproduced in 37 CFR 1.9(d), for purposes of paying reduced fees to the United States Patent and Trademark Office, in that the number of employees of the concern, including those of its affiliates, does not exceed 500 persons. For purposes of this statement, (1) the number of employees of the business concern is the average over the previous fiscal year of the concern of the persons employed on a full-time, part-time or temporary basis during each of the pay periods of the fiscal year, and (2) concerns are affiliates of each other when either, directly or indirectly, one concern controls or has the power to control the other, or a third party or parties controls or has the power to control both.

I hereby declare that rights under contract or law have been conveyed to and remain with the small business concern identified above with regard to the invention, entitled MAMMALIAN IAP GENE FAMILY, PRIMERS, PROBES, AND DETECTION METHODS by inventor(s) Robert G. Korneluk, Alexander R. MacKenzie, and Stephen Baird described in

[ ] the specification filed herewith.
[X] application serial no. 08/576,956, filed December 22, 1995.
[ ] patent no. , issued .

If the rights held by the above identified small business concern are not exclusive, each individual, concern or organization having rights to the invention is listed below* and no rights to the invention are held by any person, other than the inventor, who would not qualify as an independent inventor under 37 CFR 1.9(c) if that person made the invention, or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d), or a nonprofit organization under 37 CFR 1.9(e). NOTE: Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities. (37 CFR 1.27)

Full Name: University of Ottawa

Address: 550 Cumberland, Ottawa, Ontario, Canada K1N 6N5

[ ] INDIVIDUAL [ ] SMALL BUSINESS CONCERN [X] NONPROFIT ORGANIZATION

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent on which this verified statement is directed.

Name: Frank Gleeson

Title: President and CEO

Address: 100 Etobicoke, Ontario, Canada M9W 6J6 100 International Blvd

Signature: FM Gleeson **Date:** 21 Feb '96

PATENT
ATTORNEY DOCKET NO. 07891/003005

Certificate of Mailing

Date of Deposit September 1, 2000 Label Number: EL509049835US

I hereby certify under 37 CFR 1.10 that this correspondence is being deposited with the United States Postal Service as **"Express Mail Post Office to Addressee"** with sufficient postage on the date indicated above and is addressed to BOX PATENT APPLICATION, Director for Patents, Washington, D.C. 20231.

Luis A. Cruz
Printed name of person mailing correspondence

Signature of person mailing correspondence

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Robert G. Korneluk *et al.*	Art Unit:	Not Yet Assigned
Serial No.:	Not Yet Assigned	Examiner:	Not Yet Assigned
Filed:	September 1, 2000		
Title:	MAMMALIAN IAP GENE FAMILY, PRIMERS, PROBES, AND DETECTION METHODS		

Director for Patents
Washington, D.C. 20231

PRELIMINARY AMENDMENT

Prior to examination of the above-referenced application, kindly consider the following amendments and remarks.

Please amend the application as follows:

In the specification:

At page 1, line 5, before "The invention relates to apoptosis.", add the following:

--Cross Reference To Related Applications

This application is a continuation of U.S.S.N. 08/576,956, filed December 22, 1995, which is a continuation-in-part of U.S.S.N. 08/511,485, filed August 4, 1995, now issued as U.S. Patent No. 5,919,912.--.

At page 6, line 27, replace "IAP disease resistance gene" with --IAP gene--.

At page 18, line 15, replace "Fig. 10 is a Northern blot" with -- Figs. 10A-C are a series of Northern blots --.

At page 18, line 17, replace "Fig. 11 is a Northern blot" with -- Figs. 11A-C are a series of Northern blots --.

At page 18, line 19, replace "Fig. 12 is a Northern blot" with -- Figs. 12A-C are a series of Northern blots --.

At page 19, line 1, after "Tables 1 and 2"insert --(SEQ ID NOS: 45-92)--.

At page 24, line 23, after "MEQKLISEEDL," insert -- (SEQ ID NO: 43) --.

At page 26, line 23, replace "Embo," with -- EMBO --.

At page 27, line 3, replace "Neurobiol," with -- Neurobiol. --.

At page 27, line 27, replace "Virol," with -- Virol. --.

At page 34, line 18, replace "Cell," with -- Cell --.

At page 34, line 18, replace "Nature," with -- Nature --.

At page 36, line 8, replace "Med," with -- Med. --.

Kindly remove the sequence listing found at pages 51-88 and renumber the pages

of the claims and abstract consecutively thereafter. The enclosed amended sequence listing should be inserted at the end of the application.

In the Claims:

Cancel claims 2, 15-29, and 33-47.

Amend claims 1, 3-7, 13, 14, and 30-32 as follows.

1. (Amended) A substantially [Substantially] pure nucleic acid encoding [an IAP] a mammalian inhibitor of apoptosis protein (IAP) polypeptide, wherein said inhibitor of apoptosis protein is a protein that modulates apoptosis and comprises a ring zinc finger (RZF) domain and at least one baculovirus inhibitor of apoptosis repeat (BIR) domain.

3. (Amended) The nucleic acid of claim [2] 1, wherein said polypeptide has at least two [BIR] baculovirus inhibitor of apoptosis repeat (BIR) domains.

4. (Amended) The nucleic acid of claim 3, wherein said polypeptide has at least three [BIR] baculovirus inhibitor of apoptosis repeat (BIR) domains.

5. (Amended) The nucleic acid of claim 1, wherein said [DNA] nucleic acid contains the [xiap] X-linked inhibitor of apoptosis protein (xiap) gene.

6. (Amended) The nucleic acid of claim 1, wherein said [DNA] nucleic acid contains the [hiap2] human inhibitor of apoptosis protein 2 (hiap2) gene.

7. (Amended) The nucleic acid of claim 1, wherein said [DNA] nucleic acid contains the [hiap1] human inhibitor of apoptosis protein 1 (hiap1) gene.

DNA.

13. (Amended) A [Substantially] substantially pure [DNA] nucleic acid having the sequence of Fig. 5 (SEQ ID NO: 39), or degenerate variants thereof, and encoding the amino acid sequence of Fig. 5 (SEQ ID NO: 40).

14. (Amended) A [Substantially] substantially pure [DNA] nucleic acid having the sequence of Fig. 6 (SEQ ID NO: 41), or degenerate variants thereof, and encoding the amino acid sequence of Fig. 6 (SEQ ID NO: 42).

30. (Amended) A method of producing [an IAP] a mammalian inhibitor of apoptosis protein (IAP) polypeptide comprising:

providing a cell transformed with [DNA] nucleic acid encoding [an] a mammalian IAP polypeptide positioned for expression in said cell, said polypeptide comprising a ring zinc finger (RZF) domain;

culturing said transformed cell under conditions for expressing said [DNA] nucleic acid; and

[isolating] producing said IAP polypeptide.

31. (Amended) The method of claim 30, wherein said mammalian inhibitor of apoptosis (IAP) [IAP] polypeptide is murine human inhibitor of apoptosis protein 1 (m-HIAP1) [HIAP1].

32. (Amended) The method of claim 30, wherein said mammalian inhibitor of apoptosis (IAP) [IAP] polypeptide is murine human inhibitor of apoptosis protein 2 (m-HIAP2) [HIAP2].

Add the following new claims 48-78.

--48. A substantially pure nucleic acid that hybridizes to a probe of at least 40 nucleotides in length, said probe derived from the nucleic acid sequence of Fig. 5 (SEQ ID NO: 39), wherein said nucleic acid hybridizes to said probe under low stringency conditions, said conditions comprising washing with 2X SSC at 40°C, and wherein said nucleic acid encodes a mammalian inhibitor of apoptosis protein (IAP) polypeptide, said polypeptide comprising a ring zinc finger (RZF) domain and at least one baculovirus

inhibitor of apoptosis repeat (BIR) domain.

49. A substantially pure nucleic acid that hybridizes to a probe of at least 40 nucleotides in length, said probe derived from the nucleic acid sequence of Fig. 6 (SEQ ID NO: 41), wherein said nucleic acid hybridizes to said probe under low stringency conditions, said conditions comprising washing with 2X SSC at 40°C, and wherein said nucleic acid encodes a mammalian inhibitor of apoptosis protein (IAP) polypeptide, said polypeptide comprising a ring zinc finger (RZF) domain and at least one baculovirus inhibitor of apoptosis repeat (BIR) domain.

50. A substantially pure nucleic acid encoding a baculovirus inhibitor of apoptosis repeat (BIR) domain, said nucleic acid comprising a sequence selected from the group consisting of SEQ ID NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 66, and SEQ ID NO: 67.

51. A substantially pure nucleic acid encoding a ring zinc finger (RZF) domain, said nucleic acid comprising a sequence selected from the group consisting of SEQ ID NO: 48, SEQ ID NO: 52, SEQ ID NO: 56, SEQ ID NO: 60, SEQ ID NO: 64, and SEQ ID

NO: 68.

52. The nucleic acid of claim 1, wherein said nucleic acid encodes an X-linked inhibitor of apoptosis protein (XIAP).

53. The nucleic acid of claim 52, wherein said X-linked inhibitor of apoptosis protein (XIAP) is from a mouse.

54. The nucleic acid of claim 52, wherein said X-linked inhibitor of apoptosis protein (XIAP) is from a human.

55. The nucleic acid of claim 1, wherein said nucleic acid encodes a human inhibitor of apoptosis protein 1 (HIAP1).

56. The nucleic acid of claim 55, wherein said human inhibitor of apoptosis protein 1 (HIAP1) is from a mouse.

57. The nucleic acid of claim 55, wherein said human inhibitor of apoptosis protein 1 (HIAP1) is from a human.

58. The nucleic acid of claim 1, wherein said nucleic acid encodes a human inhibitor of apoptosis protein 2 (HIAP2).

59. The nucleic acid of claim 58, wherein said human inhibitor of apoptosis protein 2 (HIAP2) is from a mouse.

60. The nucleic acid of claim 58, wherein said human inhibitor of apoptosis protein 2 (HIAP2) is from a human.

61. The nucleic acid of claim 5, wherein said X-linked inhibitor of apoptosis protein (xiap) gene is from a mouse.

62. The nucleic acid of claim 5, wherein said X-linked inhibitor of apoptosis protein (xiap) gene is from a human.

63. The nucleic acid of claim 6, wherein said human inhibitor of apoptosis protein 2 (hiap2) gene is from a mouse.

64. The nucleic acid of claim 6, wherein said human inhibitor of apoptosis protein 2 (hiap2) gene is from a human.

65. The nucleic acid of claim 7, wherein said human inhibitor of apoptosis protein 1 (hiap1) gene is from a mouse.

66. The nucleic acid of claim 7, wherein said human inhibitor of apoptosis protein 1 (hiap1) gene is from a human.

67. A substantially pure nucleic acid having the sequence of Fig. 1 (SEQ ID NO: 3), or degenerate variants thereof, and encoding the amino acid sequence of Fig. 1 (SEQ ID NO: 4).

68. A substantially pure nucleic acid having the sequence of Fig. 2 (SEQ ID NO: 5), or degenerate variants thereof, and encoding the amino acid sequence of Fig. 2 (SEQ ID NO: 6).

69. A substantially pure nucleic acid having the sequence of Fig. 3 (SEQ ID NO: 7), or degenerate variants thereof, and encoding the amino acid sequence of Fig. 3 (SEQ ID NO: 8).

70. A substantially pure nucleic acid having the sequence of Fig. 4 (SEQ ID NO: 9), or degenerate variants thereof, and encoding the amino acid sequence of Fig. 4 (SEQ

ID NO: 10).

71. The method of claim 30, wherein said mammalian inhibitor of apoptosis protein (IAP) polypeptide is human inhibitor of apoptosis protein 1 (HIAP1).

72. The method of claim 30, wherein said mammalian inhibitor of apoptosis protein (IAP) polypeptide is human inhibitor of apoptosis protein 2 (HIAP2).

73. The method of claim 30, wherein said mammalian inhibitor of apoptosis protein (IAP) polypeptide is murine X-linked inhibitor of apoptosis protein (m-XIAP).

74. The method of claim 30, wherein said mammalian inhibitor of apoptosis protein (IAP) polypeptide is human X-linked inhibitor of apoptosis protein (XIAP).

75. A substantially pure nucleic acid that hybridizes to a probe of at least 40 nucleotides in length, said probe derived from the DNA sequence of Fig. 1 (SEQ ID NO: 3), wherein said nucleic acid hybridizes to said probe under low stringency conditions, said conditions comprising washing with 2X SSC at 40°C, and wherein said nucleic acid encodes a mammalian inhibitor of apoptosis protein (IAP) polypeptide, said polypeptide comprising a ring zinc finger (RZF) domain and at least one baculovirus inhibitor of

apoptosis repeat (BIR) domain.

76. A substantially pure nucleic acid that hybridizes to a probe of at least 40 nucleotides in length, said probe derived from the DNA sequence of Fig. 2 (SEQ ID NO: 5), wherein said nucleic acid hybridizes to said probe under low stringency conditions, said conditions comprising washing with 2X SSC at 40°C, and wherein said nucleic acid encodes a mammalian inhibitor of apoptosis protein (IAP) polypeptide, said polypeptide comprising a ring zinc finger (RZF) domain and at least one baculovirus inhibitor of apoptosis repeat (BIR) domain.

77. A substantially pure nucleic acid that hybridizes to a probe of at least 40 nucleotides in length, said probe derived from the DNA sequence of Fig. 3 (SEQ ID NO: 7), wherein said nucleic acid hybridizes to said probe under low stringency conditions, said conditions comprising washing with 2X SSC at 40°C, and wherein said nucleic acid encodes a mammalian inhibitor of apoptosis protein (IAP) polypeptide, said polypeptide comprising a ring zinc finger (RZF) domain and at least one baculovirus inhibitor of apoptosis repeat (BIR) domain.

78. A substantially pure nucleic acid that hybridizes to a probe of at least 40 nucleotides in length, said probe derived from the DNA sequence of Fig. 4 (SEQ ID NO: 9), wherein said nucleic acid hybridizes to said probe under low stringency conditions,

said conditions comprising washing with 2X SSC at 40°C, and wherein said nucleic acid encodes a mammalian inhibitor of apoptosis protein (IAP) polypeptide, said polypeptide comprising a ring zinc finger (RZF) domain and at least one baculovirus inhibitor of apoptosis repeat (BIR) domain.--

REMARKS

In general, Applicants' presently claimed invention features substantially pure nucleic acids encoding mammalian IAP polypeptides and methods of using such nucleic acids to produce such mammalian IAP polypeptides.

Support for the Amendments

The specification and drawings have been amended to comply with the requirements of 37 C.F.R. § 1.821 through 1.825. The specification has also been amended to properly refer to each individual panel of a drawing.

The specification and the claims have been amended to correct regrettable typographical errors.

Applicants have added new claims 48 and 49 to cover substantially pure DNA encoding mammalian inhibitor of apoptosis protein (IAP) polypeptides that hybridize under low stringency conditions to SEQ ID NO: 39 and SEQ ID NO: 41, respectively. Support for these new claims may be found in the specification at page 48, lines 15-20.

Applicants have added new claim 50 to cover a substantially pure DNA encoding a baculovirus inhibitor of apoptosis repeat domain that comprises the sequence of SEQ ID

NO: 45, SEQ ID NO: 46, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 50, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 54, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 66, or SEQ ID NO: 67. Support for this new claim can be found in the specification at page 19 in Table I (page 19, lines 12-20). The DNA of this claim finds use as, for example, a hybridization probe for screening libraries.

Applicants have added new claim 51 to cover a substantially pure DNA encoding a ring zinc finger domain that comprises the sequence of SEQ ID NO: 48, SEQ ID NO: 52, SEQ ID NO: 56, SEQ ID NO: 60, SEQ ID NO: 64, or SEQ ID NO: 68. Support for this new claim can be found in the specification at page 19 in Table I (page 19, lines 12-20). The DNA of this claim finds use as, for example, a hybridization probe for screening libraries.

Applicants have added new claim 52 to claim nucleic acid encoding an X-linked inhibitor of apoptosis protein (XIAP). New dependent claims 53 and 54 have been added to specifically claim nucleic acids encoding XIAP from a mouse and from a human, respectively. Support for these new claims may be found in the specification at page 21, lines 2-21, page 22, lines 8-32, and in Figs. 1 and 4.

Applicants have added new claim 55 to claim nucleic acid encoding a human inhibitor of apoptosis protein 1 (HIAP1). New dependent claims 56 and 57 have been added to specifically claim nucleic acids encoding HIAP1 from a mouse and from a human, respectively. Support for these new claims may be found in the specification at

page 21, line 22 through page 22, line 7, and in Figs. 2 and 5.

Applicants have added new claim 58 to claim nucleic acid encoding a human inhibitor of apoptosis protein 2 (HIAP2). New dependent claims 59 and 60 have been added to specifically claim nucleic acids encoding HIAP2 from a mouse and from a human, respectively. Support for these new claims may be found in the specification at page 21, line 22 through page 22, line 7, and in Figs. 3 and 6.

Applicants have added new dependent claims 61 and 62 to specifically claim nucleic acids containing the X-linked inhibitor of apoptosis (xiap) gene, where the (xiap) gene is from a mouse or a human, respectively. Support for these new claims may be found in the specification at page 21, lines 2-21, page 22, lines 8-32, and in Figs. 1 and 4.

Applicants have added new dependent claims 63 and 64 to specifically claim nucleic acids containing the human inhibitor of apoptosis 2 (hiap2) gene, where the (hiap2) gene is from a mouse or a human, respectively. Support for these new claims may be found in the specification at page 21, line 22 through page 22, line 7, and in Figs. 3 and 6.

Applicants have added new dependent claims 65 and 66 to specifically claim nucleic acids containing the human inhibitor of apoptosis 1 (hiap1) gene, where the (hiap1) gene is from a mouse or a human, respectively. Support for these new claims may be found in the specification at page 21, line 22 through page 22, line 7, and in Figs. 2 and 5.

Applicants have added new claims 67, 68, 69, and 70 to specifically claim substantially pure nucleic acids having the sequence of and encoding the amino acid sequence of Figs. 1, 2, 3, and 4, respectively. Support for these new claims may be found in the specification, for example, at page 21, line 2 through page 22, line 32, and in Figs. 1-4.

Applicants have added new dependent claims 71, 72, 73, and 74 to specifically claim methods for producing human inhibitor of apoptosis protein 1, human inhibitor of apoptosis protein 2, murine X-linked inhibitor of apoptosis protein, and human X-linked inhibitor of apoptosis protein, respectively. Support for these new claims may be found in the specification, for example, at page 5, lines 7-12; at page 21, line 2 through page 22, line 32; and in Figs. 1-4.

Applicants have added new claims 75-78 to specifically claim substantially pure nucleic acids encoding mammalian inhibitor of apoptosis protein (IAP) polypeptides that hybridize under low stringency conditions to probes derived from the DNA sequences of SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, and SEQ ID NO: 9, respectively. Support for these new claims may be found in the specification at page 48, lines 15-20, and in Figs. 1-4. No new matter is added by any of these amendments.

Sequence Listing

As required by 37 CFR 1.825(a), enclosed is an amended sequence listing consisting of 42 sheets to be inserted at the end of the application. The amendments to

the sequence listing provide each sequence in the specification with a unique SEQ ID NO, and contain no new matter. In particular, SEQ ID NOS: 69-92 have been added to include the sequences described in Table 2, found at page 20 of the specification.

As required by 37 CFR 1.825(b), also enclosed is a diskette containing a copy of the sequence listing in computer readable form including all previously submitted data with the amendments incorporated therein. The contents of the computer readable form are the same as the contents of the paper sheets.

If there are any charges or any credits, please apply them to Deposit Account No. 03-2095.

Respectfully submitted,

Date: 9/1/00

Kristina Bieker-Brady, Ph.D.
Reg. No. 39,109
James De Camp
Reg. No. 43,580

Clark & Elbing LLP
176 Federal Street
Boston, MA 02110
Telephone: 617-428-0200
Facsimile: 617-428-7045
07891.003005 Preliminary amendment xxx.wpd

Certificate of Mailing

Date of Deposit September 1, 2000 Label Number: EL509049835US

I hereby certify under 37 CFR 1.10 that this correspondence is being deposited with the United States Postal Service as **"Express Mail Post Office to Addressee"** with sufficient postage on the date indicated above and is addressed to BOX PATENT APPLICATION, *Director for Patents*, Washington, D.C. 20231.

Luis A. Cruz
Printed name of person mailing correspondence

Signature of person mailing correspondence

APPLICATION

FOR

UNITED STATES LETTERS PATENT

APPLICANT : ROBERT G. KORNELUK, ALEXANDER E. MACKENZIE, STEPHEN BAIRD, AND PETER LISTON

TITLE : MAMMALIAN IAP GENE FAMILY, PRIMERS, PROBES, AND DETECTION METHODS

MAMMALIAN IAP GENE FAMILY, PRIMERS, PROBES, AND DETECTION METHODS

The invention relates to apoptosis.

Background of the Invention

There are two general ways by which cells die. An easily recognized pathway is necrosis, a process of cell death usually resulting from severe and sudden injury. In necrosis, changes in cellular homeostasis occur with loss of membrane integrity. Dysregulation of osmotic pressure results and, as a consequence, the cells swell and finally rupture. The cellular contents are then spilled into the surrounding tissue space and, usually, an inflammation response ensues. A second form of cell death is apoptosis. This cell "suicide" pathway or programmed cell death often occurs so rapidly that in some biological systems the apoptotic process is difficult to ascertain. Indeed, it has been only in the past few years that the involvement of apoptosis in a wide spectrum of biological processes has become recognized. Apoptosis is a fundamental physiological pathway of cell death, highly conserved throughout evolution, and playing a major role in development, viral pathogenesis, cancer, autoimmune diseases and neurodegenerative disorders.

Inappropriate increases in apoptosis may cause or contribute to a variety of diseases, including AIDS, neurodegenerative diseases (e.g. Alzheimer's Disease, Parkinson's Disease, Amyotrophic Lateral Sclerosis (ALS), retinitis pigmentosa and other diseases of the retina, myelodysplastic syndrome (e.g., aplastic anemia), toxin-

induced liver disease (e.g., alcoholism) and ischemic injury (e.g., myocardial infarction, stroke, and reperfusion injury). In addition, disruption of normally occurring apoptosis has been implicated in the development of some cancers (e.g. follicular lymphoma, p53 carcinomas, and hormone dependent tumors), autoimmune disorders (e.g., lupus erythematosis and multiple sclerosis) and viral infections (e.g., herpes virus, poxvirus, and adenovirus infections).

Mature $CD4^+$ T-lymphocytes in patients with HIV-1 have been observed to respond to stimulation with mitogens or super-antigens by undergoing increased apoptosis. The great majority of these cells are not infected and similar inappropriate antigen-induced apoptosis could be very important in the destruction of this vital part of the immune system early in HIV infection.

Baculoviruses encode inhibitors of apoptosis proteins (IAPs). These proteins inhibit the apoptosis which otherwise occurs when insect cells are infected by the virus. Baculovirus IAP proteins work in a manner which is thought to be independent of other viral proteins. The baculovirus IAP genes include sequences encoding a ring zinc finger-like motif which is presumed to be involved in the direct binding of DNA.

Summary of the Invention

In general, the invention features substantially pure DNA (for example, genomic DNA, cDNA, or synthetic DNA) encoding a mammalian IAP polypeptide as defined below. In related aspects, the invention also features a vector, a cell (e.g., a mammalian, yeast or bacterial cell), and a transgenic animal or embryo thereof which includes such a substantially pure DNA encoding an IAP polypeptide.

In preferred embodiments, an IAP gene is the xiap (including human xiap and its murine homolog, m-xiap), hiap1 (including human hiap1 and m-hiap1), or the hiap2 gene (including human hiap2 and m-hiap2). In most preferred embodiments the IAP gene is a human IAP gene. In other various preferred embodiments, the cell is a transformed cell. In related aspects, the invention features a transgenic animal containing a transgene which encodes an IAP polypeptide that is expressed in or delivered to tissue normally susceptible to apoptosis.

In yet another aspect, the invention features DNA encoding fragments of IAP polypeptides including the BIR domains and the RZF domains provided herein.

In specific embodiments, the invention features DNA sequences substantially identical to the DNA sequences shown in Figs. 1-6.

In another aspect, the invention also features RNA which is encoded by the DNA described herein. Preferably, the RNA is mRNA. In another embodiment the RNA is antisense RNA.

In another aspect, the invention features a substantially pure polypeptide having a sequence substantially identical to one of the IAP amino acid sequences shown in Figures 1-6.

In a second aspect, the invention features a substantially pure DNA which includes a promoter capable of expressing the IAP gene in a cell susceptible to apoptosis. In preferred embodiments, the IAP gene is xiap (including the human or murine xiap), hiap1 (preferably the human or murine hiap1), or hiap2 (preferably the human or murine hiap2). hiap2 may be the full length gene, as shown in Fig. 3, or the truncated variant having the sequence boxed in Fig. 3 deleted.

In preferred embodiments, the promoter is the promoter native to an IAP gene. Additionally, transcriptional and translational regulatory regions are preferably native to an IAP gene.

In another aspect, the invention provides transgenic cell lines and transgenic animals. The transgenic cells of the invention are preferably cells which are susceptible to apoptosis. In preferred embodiments, the transgenic cell is a fibroblast, neuronal cell, a lymphocyte cell, or an insect cell. Most preferably, the neuron is a motor neuron and the lymphocyte is a $CD4^+$ T-cell.

In another aspect, the invention features a method of inhibiting apoptosis which involves producing a transgenic cell having a transgene encoding an IAP polypeptide wherein the transgene is integrated into the genome of the cell and is positioned for expression in the cell and wherein the IAP transgene is expressed in the cell at a level sufficient to inhibit apoptosis.

In a related aspect, the invention features a transgenic animal, preferably a mammal, more preferably a rodent, and most preferably a mouse, having either increased copies of IAP genes inserted into the genome or a knockout of an IAP gene in the genome. The transgenic animals may express an increased amount of IAP polypeptide or may express a decreased amount of an IAP polypeptide, respectively. In related embodiments, the invention provides a method of utilizing the IAP nucleic acid to engineer a knockout mutation in an IAP gene and a method of making an animal with increased expression by insertion of IAP gene into the genome.

In another aspect, the invention features a method of detecting an IAP in a cell involving: (a) contacting the IAP gene or a portion thereof greater than 9 nucleic acids,

preferably greater than 18 nucleic acids in length with a preparation of genomic DNA from the cell under hybridization conditions providing detection of DNA sequences having about 50% or greater nucleotide sequence identity to the amino acid encoding DNA sequences of hiap1, hiap2, or xiap IAP polypeptides.

In another aspect, the invention features a method of producing an IAP polypeptide which involves: (a) providing a cell transformed with DNA encoding an IAP polypeptide positioned for expression in the cell; (b) culturing the cell under conditions for expressing the DNA; and (c) isolating the IAP polypeptide. In preferred embodiments the IAP polypeptide is expressed by DNA which has a constituative or inducible promotor. In our embodiment, the promotor is a heterologous promotor.

In another aspect, the invention features substantially pure mammalian IAP polypeptide. Preferably, the polypeptide includes a greater than 50 amino acid sequence substantially identical to a greater than 50 amino acid sequence shown in any one of Figs. 1-4. Most preferably, the polypeptide is the human or murine XIAP, HIAP1, or HIAP2 polypeptide. Fragments including BIR domains and RZF-domains provided herein are also a part of the invention.

In another aspect, the invention features a recombinant mammalian polypeptide capable of modulating apoptosis wherein the polypeptide includes at least a ring zinc finger domain and a BIR domain as defined herein. In preferred embodiments, the invention features a substantially pure polypeptide and an oligonucleotide encoding said polypeptide, the polypeptide including a ring zinc finger (RZF) having the sequence:

Glu Xaa1 Xaa1 Xaa1 Xaa1 Xaa1 Xaa1 Xaa2 Xaa1 Xaa1 Xaa1 Cys Lys Xaa3 Cys Met Xaa1 Xaa1 Xaa1 Xaa1 Xaa1 Xaa3 Xaa1 Phe Xaa1 Pro Cys Gly His Xaa1 Xaa1 Xaa1 Cys Xaa1 Xaa1 Cys Ala Xaa1 Xaa1 Xaa1 Xaa1 Xaa1 Cys Pro Xaa1 Cys, wherein Xaa1 is any amino acid, Xaa2 is Glu or Asp, Xaa3 is Val or Ile (SEQ ID NO:1); and at least one BIR domain having the sequence: Xaa1 Xaa1 Xaa1 Arg Leu Xaa1 Thr Phe Xaa1 Xaa1 Trp Pro Xaa2 Xaa1 Xaa1 Xaa2 Xaa2 Xaa1 Xaa1 Xaa1 Xaa1 Leu Ala Xaa1 Ala Gly Phe Tyr Tyr Xaa1 Gly Xaa1 Xaa1 Asp Xaa1 Val Xaa1 Cys Phe Xaa1 Cys Xaa1 Xaa1 Xaa1 Xaa1 Xaa1 Xaa1 Trp Xaa1 Xaa1 Xaa1 Asp Xaa1 Xaa1 Xaa1 Xaa1 Xaa1 His Xaa1 Xaa1 Xaa1 Xaa1 Pro Xaa1 Cys Xaa1 Phe Val, wherein Xaa1 may be any amino acid and Xaa2 may be any amino acid or may be absent (SEQ ID NO:2).

In various preferred embodiments the protein has at least two or, more preferably at least three BIR domains, the RZF domain has one of the IAP sequences shown in Fig. 6, and the BIR domains are comprised of BIR domains shown in Fig. 5. In other preferred embodiments the BIR domains are at the amino terminal end of the protein relative to the RZF domain, which is at or near the carboxy terminus of the polypeptide.

In another aspect, the invention features an IAP gene isolated according to the method involving: (a) providing a sample of DNA; (b) providing a pair of oligonucleotides having sequence homology to a conserved region of an IAP disease-resistance gene; (c) combining the pair of oligonucleotides with the cell DNA sample under conditions suitable for polymerase chain reaction-mediated DNA amplification; and (d) isolating the amplified IAP gene or fragment thereof.

In preferred embodiments, the amplification is carried out using a reverse-transcription polymerase chain reaction, for example, the RACE method.

In another aspect, the invention features an IAP gene isolated according to the method involving: (a) providing a preparation of DNA; (b) providing a detectably-labelled DNA sequence having homology to a conserved region of an IAP gene; (c) contacting the preparation of DNA with the detectably-labelled DNA sequence under hybridization conditions providing detection of genes having 50% or greater nucleotide sequence identity; and (d) identifying an IAP gene by its association with the detectable label.

In another aspect, the invention features an IAP gene isolated according to the method involving: (a) providing a cell sample; (b) introducing by transformation into the cell sample a candidate IAP gene; (c) expressing the candidate IAP gene within the cell sample; and (d) determining whether the cell sample exhibits an altered apoptotic response, whereby a response identifies an IAP gene.

In another aspect, the invention features a method of identifying an IAP gene in a cell, involving: (a) providing a preparation of cellular DNA (for example, from the human genome or a cDNA library (such as a cDNA library isolated from a cell type which undergoes apoptosis); (b) providing a detectably-labelled DNA sequence (for example, prepared by the methods of the invention) having homology to a conserved region of an IAP gene; (c) contacting the preparation of cellular DNA with the detectably-labelled DNA sequence under hybridization conditions providing detection of genes having 50% nucleotide or greater sequence identity; and (d) identifying an IAP gene by its association with the detectable label.

In another aspect, the invention features a method of isolating an IAP gene from a recombinant library, involving: (a) providing a recombinant library; (b) contacting the library with a detectably-labelled gene fragment produced according to the PCR method of the invention under hybridization conditions providing detection of genes having 50% or greater nucleotide sequence identity; and (c) isolating an IAP gene by its association with the detectable label.

In another aspect, the invention features a method of identifying an IAP gene involving: (a) providing a cell tissue sample; (b) introducing by transformation into the cell sample a candidate IAP gene; (c) expressing the candidate IAP gene within the cell sample; and (d) determining whether the cell sample exhibits inhibition of apoptosis, whereby a change in (i.e. modulation of) apoptosis identifies an IAP gene.

Preferably, the cell sample is a cell type which may be assayed for apoptosis (e.g., lymphocytes, T-cells and B-cells, neuronal cells, baculovirus infected insect cells and fibroblast cells); the candidate IAP gene is obtained from a cDNA expression library; and the apoptosis response is the inhibition of apoptosis.

In another aspect, the invention features a method of inhibiting apoptosis in a mammal wherein the method includes: (a) providing DNA encoding at least one IAP polypeptide to a cell which is susceptible to apoptosis; wherein the DNA is integrated into the genome of the cell and is positioned for expression in the cell; and the IAP gene is under the control of regulatory sequences suitable for controlled expression of the gene(s); wherein the IAP transgene is expressed at a level sufficient to inhibit apoptosis relative to a cell lacking the IAP transgene. It

will be appreciated that IAP polypeptides also may be administered directly to inhibit any undesirable apoptosis.

In a related aspect, the invention features a method of inhibiting apoptosis wherein the method involves: (a) producing a cell having integrated in the genome a transgene containing the IAP gene under the control of a promoter providing constitutive expression of the IAP gene.

In yet another related aspect, the invention features a method of inhibiting apoptosis wherein the method involves: (a) producing a cell having integrated in the genome a transgene containing the IAP gene under the control of a promoter providing controllable expression of the IAP gene; and (b) regulating the environment of the cell so that the IAP transgene is controllably expressed in the cell. In preferred embodiments, the IAP gene is expressed using a tissue-specific or cell type-specific promoter, or by a promoter that is activated by the introduction of an external signal or agent, such as a chemical signal or agent. In preferred embodiments the cell is a lymphocyte or B-cell, a neuronal cell, or a fibroblast. In other embodiments the cell is a cell in an HIV infected human, or a mammal with a neurodegenerative disease, ischemia, toxin induced liver disease, or a myelodysplastic syndrome.

In a related aspect, the invention provides a method of inhibiting apoptosis in a mammal by providing an apoptosis-inhibiting amount of IAP polypeptide.

In another aspect, the invention features a purified antibody which binds specifically to an IAP family protein. Such an antibody may be used in any standard immunodetection method for the identification of an IAP polypeptide. Preferably, the antibody binds specifically to xiap, hiap1 or hiap2. In various embodiments the antibody may react with other IAP polypeptides or may be specific for one or a

few IAP polypeptides. The antibody may be a monoclonal polyclonal antibody. Preferably, the antibody reacts specifically with only one of the IAP polypeptides, for example, reacts with murine and human xiap, but not with hiap1 or hiap2 from mammalian species.

In another aspect, the invention features a method of identifying a compound which modulates apoptosis. The method includes (a) providing a cell expressing an IAP polypeptide; and (b) contracting the cell with a candidate compound, and monitoring the expression of an IAP gene. An alteration in the level of expression of the IAP gene indicates the presence of a compound which modulates apoptosis. The compound may be an inhibitor or an enhancer of apoptosis. In various preferred embodiments, the cell is a fibroblast, a neuronal cell, a lymphocyte (T-cell or B-cell), or an insect cell; the polypeptide expression being monitored is XIAP, HIAP1, or HIAP2 (e.g., human or murine).

In a related aspect, the invention features methods of detecting compounds which modulate apoptosis using the interaction trap technology and IAP polypeptides or fragments thereof as a component of the bait. In preferred embodiments, the compound being tested as a modulator of apoptosis is also a polypeptide.

In another aspect, the invention features a method for diagnosing a cell proliferation disease, or an increased liklihood of such a disease, using an IAP nucleic acid probe or antibody. Preferably, the disease is a cancer. Most preferably, the disease is selected from the group consisting of promyelocytic leukemia, a Hela-type carcinoma, chronic myelogenous leukemia (preferably using xiap or hiap2 related probes), lymphoblastic leukemia (preferably using a xiap related probe), Burkitt's lymphoma (preferably using an hiap1 related probe), colorectal adenocarcinoma, lung

carcinoma, and melanoma (preferably using a xiap probe). Preferably, a diagnosis is indicated by a 2-fold increase in expression or activity, more preferably, at least a 10-fold increase in expression or activity.

By "IAP gene" is meant a gene encoding a polypeptide having at least one BIR domain and a ring zinc finger domain which is capable of modulating (inhibiting or enhancing) apoptosis in a cell or tissue when provided by other intracellular or extracellular delivery methods. In preferred embodiments the IAP gene is a gene having about 50% or greater nucleotide sequence identity to at least one of the IAP amino acid encoding sequences of Figs. 1-4 or portions thereof. Preferably, the region of sequence over which identity is measured is a region encoding at least one BIR domain and a ring zinc finger domain. Mammalian IAP genes include nucleotide sequences isolated from any mammalian source. Preferably, the mammal is a human.

By an "IAP gene" is also meant any member of the family of apoptosis inhibitory genes characterized by their ability to modulate apoptosis and having at least 20%, preferably 30%, and most preferably 50% amino acid sequence identity to at least one of the conserved regions of one of the IAP members described herein (i.e., either the BIR or ring zinc finger domains from the human or murine xiap, hiap1 and hiap2). Representative members of the IAP gene family include, without limitation, the human and murine xiap, hiap1, and hiap2 genes. By "IAP protein" is meant a polypeptide encoded by an IAP gene.

By "BIR domain" is meant a domain having the amino acid sequence of the consensus sequence: Xaa1 Xaa1 Xaa1 Arg Leu Xaa1 Thr Phe Xaa1 Xaa1 Trp Pro Xaa2 Xaa1 Xaa1 Xaa2 Xaa2 Xaa1 Xaa1 Xaa1 Xaa1 Leu Ala Xaa1 Ala Gly Phe Tyr Tyr Xaa1 Gly Xaa1 Xaa1 Asp Xaa1 Val Xaa1 Cys Phe Xaa1 Cys Xaa1 Xaa1

Xaa1 Xaa1 Xaa1 Xaa1 Trp Xaa1 Xaa1 Xaa1 Asp Xaa1 Xaa1 Xaa1 Xaa1 Xaa1 His Xaa1 Xaa1 Xaa1 Xaa1 Pro Xaa1 Cys Xaa1 Phe Val, wherein Xaa1 is any amino acid and Xaa2 is any amino acid or is absent (SEQ ID NO:2). Preferably, the sequence is substantially identical to one of the BIR domain sequences provided for xiap, hiap1, hiap2 herein.

By "ring zinc finger" or "RZF" is meant a domain having the amino acid sequence of the consensus sequence: Glu Xaa1 Xaa1 Xaa1 Xaa1 Xaa1 Xaa1 Xaa2 Xaa1 Xaa1 Xaa1 Cys Lys Xaa3 Cys Met Xaa1 Xaa1 Xaa1 Xaa1 Xaa1 Xaa3 Xaa1 Phe Xaa1 Pro Cys Gly His Xaa1 Xaa1 Xaa1 Cys Xaa1 Xaa1 Cys Ala Xaa1 Xaa1 Xaa1 Xaa1 Xaa1 Cys Pro Xaa1 Cys, wherein Xaa1 is any amino acid, Xaa2 is Glu or Asp, and Xaa3 is Val or Ile (SEQ ID NO:1). Preferably, the sequence is substantially identical to the RZF domains provided herein for the human or murine xiap, hiap1, or hiap2.

By "modulating apoptosis" or "altering apoptosis" is meant increasing or decreasing the number of cells which undergo apoptosis in a given cell population. Preferably, the cell population is selected from a group including T-cells, neuronal cells, fibroblasts, or any other cell line known to undergo apoptosis in a laboratory setting (e.g., the baculovirus infected insect cells). It will be appreciated that the degree of modulation provided by an IAP or modulating compound in a given assay will vary, but that one skilled in the art can determine the statistically significant change in the level of apoptosis which identifies an IAP or a compound which modulates an IAP.

By "inhibiting apoptosis" is meant any decrease in the number of cells which undergo apoptosis relative to an untreated control. Preferably, the decrease is at least 25%, more preferably the decrease is 50%, and most preferably the decrease is at least one-fold.

By "polypeptide" is meant any chain of amino acids, regardless of length or post-translational modification (e.g., glycosylation or phosphorylation).

By "substantially identical" is meant a polypeptide or nucleic acid exhibiting at least 50%, preferably 85%, more preferably 90%, and most preferably 95% homology to a reference amino acid or nucleic acid sequence. For polypeptides, the length of comparison sequences will generally be at least 16 amino acids, preferably at least 20 amino acids, more preferably at least 25 amino acids, and most preferably 35 amino acids. For nucleic acids, the length of comparison sequences will generally be at least 50 nucleotides, preferably at least 60 nucleotides, more preferably at least 75 nucleotides, and most preferably 110 nucleotides.

Sequence identity is typically measured using sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705). Such software matches similar sequences by assigning degrees of homology to various substitutions, deletions, and other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine, valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine.

By a "substantially pure polypeptide" is meant an IAP polypeptide which has been separated from components which naturally accompany it. Typically, the polypeptide is substantially pure when it is at least 60%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the

preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight, IAP polypeptide. A substantially pure IAP polypeptide may be obtained, for example, by extraction from a natural source (e.g., a fibroblast, neuronal cell, or lymphocyte cell); by expression of a recombinant nucleic acid encoding an IAP polypeptide; or by chemically synthesizing the protein. Purity can be measured by any appropriate method, e.g., those described in column chromatography, polyacrylamide gel electrophoresis, or by HPLC analysis.

A protein is substantially free of naturally associated components when it is separated from those contaminants which accompany it in its natural state. Thus, a protein which is chemically synthesized or produced in a cellular system different from the cell from which it naturally originates will be substantially free from its naturally associated components. Accordingly, substantially pure polypeptides include those derived from eukaryotic organisms but synthesized in *E. coli* or other prokaryotes.

By "substantially pure DNA" is meant DNA that is free of the genes which, in the naturally-occurring genome of the organism from which the DNA of the invention is derived, flank the gene. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote; or which exists as a separate molecule (e.g., a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence.

By "transformed cell" is meant a cell into which (or into an ancestor of which) has been introduced, by means of

recombinant DNA techniques, a DNA molecule encoding (as used herein) an IAP polypeptide.

By "transgene" is meant any piece of DNA which is inserted by artifice into a cell, and becomes part of the genome of the organism which develops from that cell. Such a transgene may include a gene which is partly or entirely heterologous (i.e., foreign) to the transgenic organism, or may represent a gene homologous to an endogenous gene of the organism.

By "transgenic" is meant any cell which includes a DNA sequence which is inserted by artifice into a cell and becomes part of the genome of the organism which develops from that cell. As used herein, the transgenic organisms are generally transgenic mammalian (e.g., rodents such as rats or mice) and the DNA (transgene) is inserted by artifice into the nuclear genome.

By "transformation" is meant any method for introducing foreign molecules into a cell. Lipofection, calcium phosphate precipitation, retroviral deliver, electroporation and biolistic transformation are just a few of the teachings which may be used. For example, Biolistic transformation is a method for introducing foreign molecules into a cell using velocity driven microprojectiles such as tungsten or gold particles. Such velocity-driven methods originate from pressure bursts which include, but are not limited to, helium-driven, air-driven, and gunpowder-driven techniques. Biolistic transformation may be applied to the transformation or transfection of a wide variety of cell types and intact tissues including, without limitation, intracellular organelles (e.g., and mitochondria and chloroplasts), bacteria, yeast, fungi, algae, animal tissue, and cultured cells.

By "positioned for expression" is meant that the DNA molecule is positioned adjacent to a DNA sequence which directs transcription and translation of the sequence (i.e., facilitates the production of, e.g., an IAP polypeptide, a recombinant protein or a RNA molecule).

By "reporter gene" is meant a gene whose expression may be assayed; such genes include, without limitation, ß-glucuronidase (GUS), luciferase, chloramphenicol transacetylase (CAT), and ß-galactosidase.

By "promoter" is meant minimal sequence sufficient to direct transcription. Also included in the invention are those promoter elements which are sufficient to render promoter-dependent gene expression controllable for cell-type specific, tissue-specific or inducible by external signals or agents; such elements may be located in the 5′ or 3′ regions of the native gene.

By "operably linked" is meant that a gene and a regulatory sequence(s) are connected in such a way as to permit gene expression when the appropriate molecules (e.g., transcriptional activator proteins) are bound to the regulatory sequence(s).

By "conserved region" is meant any stretch of six or more contiguous amino acids exhibiting at least 30%, preferably 50%, and most preferably 70% amino acid sequence identity between two or more of the IAP family members, (e.g., between human HIAP1, HIAP2, and XIAP). Examples of preferred conserved regions are shown (as boxed or designated sequences) in Figures 5-7 and Tables 1 and 2, and include, without limitation, BIR domains and ring zinc finger domains.

By "detectably-labelled" is meant any means for marking and identifying the presence of a molecule, e.g., an oligonucleotide probe or primer, a gene or fragment thereof,

or a cDNA molecule. Methods for detectably-labelling a molecule are well known in the art and include, without limitation, radioactive labelling (e.g., with an isotope such as ^{32}P or ^{35}S) and nonradioactive labelling (e.g., chemiluminescent labelling, e.g., fluorescein labelling).

By "purified antibody" is meant antibody which is at least 60%, by weight, free from proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably 90%, and most preferably at least 99%, by weight, antibody, e.g., an IAP specific antibody. A purified antibody may be obtained, for example, by affinity chromatography using recombinantly-produced protein or conserved motif peptides and standard techniques.

By "specifically binds" is meant an antibody which recognizes and binds a protein but which does not substantially recognize and bind other molecules in a sample, e.g., a biological sample, which naturally includes protein.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

Detailed Description

The drawings will first be described.

Drawings

Fig. 1 is the human xiap cDNA sequence and the XIAP polypeptide sequence (SEQ ID NOS:3, 4).

Fig. 2 is the human hiap1 cDNA sequence and the HIAP1 polypeptide sequence (SEQ ID NOS:5, 6).

Fig. 3 is the human hiap2 cDNA sequence and the HIAP2 polypeptide sequence (SEQ ID NOS:7, 8). The sequence absent in the hiap2-G variant is boxed.

Fig. 4 is the murine xiap cDNA sequence and encoded murine XIAP polypeptide sequence (SEQ ID NOS:9, 10).

Fig. 5 is the murine hiap1 cDNA sequence and the encoded murine HIAP1 polypeptide sequence (SEQ ID NOS:39, 40).

Fig. 6 is the murine hiap2 cDNA sequence and the encoded murine HIAP2 polypeptide SEQ ID NOS:41, 42).

Fig. 7 shows the alignment of the BIR domains of IAP proteins (SEQ ID NOS: 11 and 14-31).

Fig. 8 is the alignment of human IAP polypeptides with diap, cp-iap, and the consensus sequence (SEQ ID NOS:4, 6, 8, 10, 12, and 13).

Fig. 9 shows the alignment of the Ring Zinc Finger domains of IAP proteins (SEQ ID NOS: 32-38).

Fig. 10 is a Northern blot showing human hiap1 and hiap2 mRNA expression in human tissues.

Fig. 11 is a Northern blot showing human hiap2 mRNA expression in human tissues.

Fig. 12 is a Northern blot showing human xiap mRNA expression in human tissues.

Fig. 13A and 13B are agarose gels showing apoptic DNA ladders and RT PCR products using hiap1 and hiap2 specific probes in HIV infected T cells.

Fig. 14A - 14D are graphs showing apoptosis suppression by XIAP, HIAP1, HIAP2, bcl-2m, smn and 6-myc.

I. IAP Polypeptides and Genes Encoding IAP polypeptides

We have discovered a new class of mammalian proteins which modulate apoptosis (IAPs) and the genes which encode these proteins. The IAP proteins are characterized by the presence of a ring zinc finger (RZF) domain (Fig. 9) and at least one BIR domain as defined by the boxed consensus sequences in Figs. 7 and 8 and by the sequence domains

provided in Tables 1 and 2. As examples of the IAP proteins we provide the cDNA sequences and amino acid sequences for these new human and murine apoptosis inhibitors, HIAP1, HIAP2, and XIAP. Additional members of the mammalian IAP family (including homologs from other species and mutant sequences) may be isolated using standard cloning techniques and the conserved amino acid sequences, primers and probes provided herein and known in the art.

This application is related to U.S. Serial Number 08/511,485, filed August 4, 1995. U.S.S.N 08/511,485 is hereby incorporated by reference.

TABLE 1

NUCLEOTIDE POSITION OF CONSERVED DOMAINS*

	BIR-1	**BIR-2**	**BIR-3**	**Ring Zinc Finger**
h-xiap	109 - 312	520 - 723	826 - 1023	1348-1485
m-xiap	202 - 405	613 - 816	916 - 1113	1438-1575
h-hiap1	273 - 476	693 - 893	951 - 1154	1824-1961
m-hiap1	251 - 453	670 - 870	928 - 1131	1795-1932
h-hiap2	373 - 576	787 - 987	1042-1245	1915-2052
m-hiap2	215 - 418	608 - 808	863 - 1066	1763-1876

* Positions indicate correspond to those shown in Figs. 1-4.

TABLE 2

AMINO ACID POSITION OF CONSERVED DOMAINS*

	BIR-1	BIR-2	BIR-3	Ring Zinc Finger
h-Xiap	26 - 93	163 - 230	265 - 330	439 - 484
m-Xiap	26 - 93	163 - 230	264 - 329	438 - 483
h-Hiap1	29 - 96	169 - 235	255 - 322	546 - 591
m-Hiap1	29 - 96	169 - 235	255 - 322	544 - 589
h-Hiap2	46 - 113	184 - 250	269 - 336	560 - 605
m-Hiap2	25 - 92	156 - 222	241 - 308	541 - 578

* Positions indicate correspond to those shown in Figs. 1-4.

Recognition of this mammalian IAP family has provided emergent patterns of protein structure. Recognition of these patterns has also allowed us assign the function of a modulator of apoptosis to a drosophila gene product of previously unknown function (Genbank Accession Number M96581). The amino acid sequence of this protein, termed diap, is shown in Fig. 8 for comparison.

The IAP proteins may be used to inhibit the apoptosis which occurs as part of disease or disorder processes. For example, IAP polypeptides or nucleic acid encoding IAP polypeptides may be administered for the treatment of or prevention of apoptosis which occurs as a part of AIDS, neurodegenerative diseases, ischemic injury, toxin-induced liver disease and myelodysplastic syndromes. Nucleic acid encoding the IAP polypeptide may also be provided to inhibit apoptosis.

II. Cloning of IAP Genes

A. XIAP

Our search for human genes potentially involved in apoptosis has resulted in the identification of an x-linked sequence tag site (STS) in the GenBank which demonstrated strong homology with the conserved RZF domain of CpIAP and OpIAP, the two baculovirus genes known to inhibit apoptosis (Clem et al., Mol. Cell Biol., 14:5212-5222, (1994); and Birnbaum et al, J. Virol. 68:2521-8, (1994)). Screening a human fetal brain ZapII cDNA library (Stratagene, La Jolla, CA) with this STS resulted in the identification and cloning of xiap (for X-linked Inhibitor of apoptosis protein gene). The human gene has a 1.7 kb coding sequence that includes three BIR (baculovirus inhibitor of apoptosis repeat (Crook et al., J. Virol. 67:2168-74, (1993), Clem et al., Science 254:1388-90, (1991); and Birnbaum et al., J. Virol., 68:2521-8, (1994)) domains and a zinc finger. Northern analysis with xiap reveals a greater than 7kb message expressed in different tissues particularly liver and kidney (Fig. 12). The large size of the transcript reflects large 5′ and 3′ untranslated regions.

B. HUMAN HIAP1 and HIAP2

The hiap1 and hiap2 genes were cloned by screening a human liver library (Stratagene) with a probe including the whole xiap coding region at low stringency (40°C wash, 2xssc, 10% SDS) (Figs. 2 and 3). hiap1 and hiap2 were also independently detected using a probe derived from a expressed sequence tag (EST) (GenBank Accession No. T96284) which includes a portion of a BIR domain. This EST was originally isolated by the PCR amplification of a cDNA library using the EST-specific primers. The derived probe

was then used to screen the human liver cDNA library for full length hiap coding sequences. We have subsequently detected a third DNA which includes the hiap2 sequence which appears to lack one exon, presumably due to alternative mRNA splicing (see boxed region in Fig. 3). Figures 8 and 9 show hiap1 and hiap2 expression in human tissues as assayed by Northern Analysis.

C. M-XIAP

Screening of a mouse embryo λgt11 cDNA library (Clonetech, Palo Alto, CA) and a mouse FIX II genomic library with a *xiap* cDNA clones probe has resulted in the identification of 14 positive cDNA and two hybridizing genomic clones. A cDNA contig spanning 8.0 kb was constructed using 12 overlapping mouse clones. DNA sequencing revealed a coding sequence of about 1.7 kb. The mouse gene called *m-xiap* (for mouse x-linked inhibitor of apoptosis protein gene) shows striking amino acid homology with xiap at and around the initiation methionine, the stop codon, the three BIR domains and the zinc finger domain. As with the human gene, the mouse homologue contains large 5′ and 3′ UTRs predicted to result in a transcript as large as 7-8 kb.

Sequencing and restriction mapping of m-xiap can be used to further delineate the structure and genomic organization of m-xiap. Southern blot analysis and inverse PCR technique (Groden et al., Cell 66:589-600 (1991) can be employed to map exons and sequence exon-intron boundaries.

Antisera can be raised against a m-xiap fusion protein expressed in *Escherichia coli* using a bacterial expression system. The resulting antisera can be used along with Northern blot analysis to analyze the spatial and temporal expression of *m-xiap* in the mouse.

D. M-HIAP1 and M-HIAP2

The murine homologs to hiap1 and hiap2 were cloned and sequenced in the same general manner as m-xiap using the human hiap1 and hiap2 sequences as probes. Cloning of m-hiap1 and m-hiap2 provide further demonstrations of the case with which homologs from different species may be detected and obtained using the techniques provided herein and those generally known to one skilled in the art of molecular biology.

III. Cloning of Additional IAP Genes

Low stringency Southern blot hybridization of human genomic DNA using probes specific for xiap, hiap1 and hiap2 show bands which correspond to the other known human IAP sequences. In addition, these probes detect sequences which do not correspond to known IAP sequences. This result indicates that additional IAP sequences may be readily identified using low stringency hybridization. Examples of murine and human xiap, hiap1, and hiap2 specific primers which may be used to clone additional genes by RT PCR are shown in Table 5. Standard techniques including PCR and hybridization may be used to clone homologs and additional genes.

IV. Characterization of IAP Apoptosis Modulating Activity

The apoptosis inhibiting capability of IAPs can be defined in an *in vitro* system know to detect alterations in apoptosis. Mammalian expression constructs carrying IAPs and their truncated forms can be introduced into various cell lines such as CHO, HIH 3T3, HL60, Rat-1, or Jurkart cells, for example. In addition, SF21 insect cells may be used in which case the IAP gene is preferentially expressed using an insect heat shock promotor. Apoptosis will then be

induced in transfected cells and controls employing standard methodologies (e.g. serum withdrawal and staurosporine). A survival index (ratio of surviving transfected cells to surviving control cells) will indicate the strength of each IAP construct in inhibiting apoptosis. These experiments can confirm the presence of apoptosis inhibiting or enhancing activity and, can help to determine the minimal functional region of an IAP. These methods may also be used in combination with compounds to identify compounds which modulate apoptosis via their effect on IAP expression.

Figs. 14A - 14D show specific examples of apoptosis suppression assays. Fig. 14A shows CHO survival following serum withdrawal. CHO cells were transfected via Lipofectace with 2 μg of each of the following recombinant plasmids; pCDNA3-6myc-hiap-1, pCDNA3-6myc-hiap-2, pCDNA3-6myc-xiap, pCDNA3-6myc, pCDNA3-HA-smn, and pCDNA3-bcl-2. Oligonucleotide primers were synthesized to allow PCR amplification and cloning of the xiap, hiap-1 and hiap-2. Oligonucleotide primers were synthesized to allow PCR amplification and cloning of the xiap, hiap-1, and hiap-2 ORFs in pCDNA3 (Invitrogen). Each construct was modified to incorporate a synthetic myc tag encoding six repeats of the peptide sequence MEQKLISEEDL allowing detection of myc-IAP fusion proteins via monoclonal anti-myc antiserum (Egan, et al., Nature 363:45-51, 1993). Triplicate samples of cell lines in 24 well dishes were washed 5 times with serum free media and maintained in serum free conditions during the course of the experiment. Trypan blue exclusion counting of viable cells utilizing a hemocytometer was performed on samples at time zero, 24 hrs., 48 hrs., and 72 hrs., post serum withdrawal. Survival was calculated as a percentage of initial numbers. Numbers represent the average of three separate experiments performed in triplicate, +/- average

deviation. Fig. 14B shows survival of CHO transfected cell lines following exposure to menadione. Cell lines were plated in 24 well dishes, allowed to grow overnight, then exposed for 1.5 hrs. to [20mM] menadione (Sigma). Triplicate samples were harvested at the time of exposure and at 24 hrs. post exposure and assessed by trypan blue exclusion for survival. Data represents the average of three independent experiments, +/- average deviation. Fig. 14C shows survival of Rat-1 cells following staurosporine exposure. Rat-1 cells were transfected with the plasmids listed in a), with selection in [800 mg/ml] G418 media for two weeks. Cell lines were assessed for resistance to [1μM]staurosporine induced apoptosis for 5 hrs. Viable cell counts were obtained 24 hrs. post exposure via trypan blue exclusion counting of samples prepared in triplicate. Numbers represent the average of two independent experiments, +/- average deviation. Fig. 14D shows Rat-1 cell lines were tested for resistance to [10 mM] menadione for 1.5 hrs., then counted at 18 hrs. post exposure. Numbers represent the average of three experiments performed in triplicate, +/- average deviation.

Specific examples of apoptosis assays are also provided in the following references:

Lymphocyte: C.J. Li et al., "Induction of apoptosis in uninfected lymphocytes by HIV-1 Tat protein", Science, 268:429-431 (1995); D. Gibellini et al., "Tat-expressing Jurkat cells show an increased resistance to different apoptotic stimuli, including acute human immunodeficiency virus-type 1 (HIV-1) infection", Br. J. Haematol. 89:24-33, (1995); S.J. Martin et al., "HIV-1 infection of human CD4+ T cells in vitro. Differential induction of apoptosis in

these cells." J. Immunol. 152:330-42, (1994); C. Terai et al., "Apoptosis as a mechanism of cell death in cultured T lymphoblasts acutely infected with HIV-1", J. Clin Invest., 87:1710-5, (1991); J. Dhein et al., "Autocrine T-cell suicide mediated by APO-1/(Fas/CD95)", Nature 373:438-441, (1995); P.D. Katsikis et al., "Fas antigen stimulation induces marked apoptosis of T lymphocytes in human immunodeficiency virus-infected individuals", J. Exp. Med. 1815:2029-2036, (1995); Michael O. Westendorp et al., Sensitization of T cells to CD95-mediated apoptosis by HIV-1 Tat and gp120", Nature, 375:497, (1995); DeRossi et al., Virology 198:234-44, (1994).

Fibroblasts: H. Vossbeck et al., "Direct transforming activity of TGF-beta on rat fibroblasts", Int. J. Cancer, 61:92-97, (1995); S. Goruppi et al., "Dissection of c-myc domains involved in S phase induction of NIH3T3 fibroblasts", Oncogene, 9:1537-44, (1994); A. Fernandez et al., "Differential sensitivity of normal and Ha-ras-transformed C3H mouse embryo fibroblasts to tumor necrosis factor: induction of bcl-2, c-myc, and manganese superoxide dismutase in resistant cells", Oncogene, 9:2009-17, (1994); E. A. Harrington et al., "c-Myc-induced apoptosis in fibroblasts in inhibited by specific cytokines", Embo J., 13:3286-3295, (1994); N. Itoh et al., "A novel protein domain required for apoptosis. Mutational analysis of human Fas antigen", J. Biol. Chem., 268:10932-7, (1993).

Neuronal Cells: G. Melino et al., "Tissue transglutaminase and apoptosis: sense and antisense transfection studies with human neuroblastoma cells", Mol. Cell. Biol., 14:6584-6596, (1994); D. M. Rosenbaum et al., "Evidence for hypoxia-induced, programmed cell death of

cultured neurons", Ann. Neurol., 36:864-870, (1994); N. Sato et al., "Neuronal differentiation of PC12 cells as a result of prevention of cell death by bcl-2", J. Neurobiol, 25:1227-1234, (1994); G. Ferrari et al., "N-acetylcysteine (D- and L-stereoisomers) prevents apoptotic death of neuronal cells", J. Neurosci., 1516:2857-2866, (1995); A. K. Talley et al., "Tumor necrosis factor alpha-induced apoptosis in human neuronal cells: protection by the antioxidant N-acetylcysteine and the genes bcl-2 and crmA", Mol. Cell Biol., 1585:2359-2366, (1995); A. K. Talley et al., "Tumor Necrosis Factor Alpha-Induced Apoptosis in Human Neuronal Cells: Protection by the Antioxidant N-Acetylcysteine and the Genes bcl-2 and crmA", Mol. and Cell. Biol., 15:2359-2366, (1995); G. Walkinshaw et al., "Induction of apoptosis in catecholaminergic PC12 cells by L-DOPA. Implications for the treatment of Parkinson's disease.", J. Clin. Invest. 95:2458-2464, (1995).

Insect Cells: R. J. Clem et al., "Prevention of apoptosis by a baculovirus gene during infection of insect cells", Science, 254:1388-90, (1991); N. E. Crook et al., "An apoptosis-inhibiting baculovirus gene with a zinc finger-like motif", J. Virol., 67:2168-74, (1993); S. Rabizadeh et al., "Expression of the baculovirus p35 gene inhibits mammalian neural cell death", J. Neurochem., 61:2318-21, (1993); M. J. Birnbaum et al., "An apoptosis-inhibiting gene from a nuclear polyhidrosis virus encoding a polypeptide with Cys/His sequence motifs", J. Virol, 68:2521-8, (1994); R. J. Clem et al., "Control of programmed cell death by the baculovirus genes p35 and iap", Mol. Cell. Biol., 14:5212-5222, (1994).

V. Construction of a Transgenic Animal

Characterization of IAPs can provide information that allows for the development of an IAP knockout animal model, preferably mammal, most preferably a mouse, by homologous recombination. Similarily, an IAP overproducing animal may be produced by means of DNA sequence integration into the genome.

A replacement type targeting vector to create a knockout can be constructed using an isogenic genomic clone from a mouse strain, e.g. 129/Sv (Strategene LaJolla, CA). The targeting vector will be introduced into a J1 line of embryonic stem (ES) cells by electroporation to generate ES cell lines that carry a profoundly truncated form of an IAP. To generate chimeric founder mice, the targeted cell lines will be injected into a mouse blastula stage embryo. Heterozygote offspring will be interbred to homozygosity. Knockout mice may be constructed as a means of screening in vivo for therapeutic compounds which modulate apoptosis.

Animals having enhanced IAP expression may also be constructed using standard transgenic technologies.

VI. IAP Protein Expression

IAP genes may be expressed in both prokaryotic and eukaryotic cell types. For those IAP's which increase apoptosis it may be desirable to express the protein under control of an inducible promotor for the purposes of protein production.

In general, IAP proteins according to the invention may be produced by transformation of a suitable host cell with all or part of a IAP-encoding cDNA fragment (e.g., the cDNA described above) in a suitable expression vehicle.

Those skilled in the field of molecular biology will understand that any of a wide variety of expression systems

may be used to provide the recombinant protein. The precise host cell used is not critical to the invention. The IAP protein may be produced in a prokaryotic host (e.g., E. coli) or in a eukaryotic host (e.g., Saccharomyces cerevisiae, insect cells, e.g., Sf21 cells, or mammalian cells, e.g., COS 1, NIH 3T3, or HeLa cells). Such cells are available from a wide range of sources (e.g., the American Type Culture Collection, Rockland, MD; also, see, e.g., Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1994). The method of transformation or transfection and the choice of expression vehicle will depend on the host system selected. Transformation and transfection methods are described, e.g., in Ausubel et al., (supra); expression vehicles may be chosen from those provided, e.g., in Cloning Vectors: A Laboratory Manual (P.H. Pouwels et al., 1985, Supp. 1987).

One preferred expression system is the baculovirus system (using, for example, the vector pBacPAK9) available from Clontech (Palo Alto, CA). If desired, this system may be used in conjunction with other protein expression techniques, for example, the myc tag approach described by Evan et al. (Mol. Cell Biol. 5:3610-3616, 1985).

Alternatively, a IAP protein is produced by a stably-transfected mammalian cell line. A number of vectors suitable for stable transfection of mammalian cells are available to the public, e.g., see Pouwels et al. (supra); methods for constructing such cell lines are also publicly available, e.g., in Ausubel et al. (supra). In one example, cDNA encoding the IAP protein is cloned into an expression vector which includes the dihydrofolate reductase (DHFR) gene. Integration of the plasmid and, therefore, the IAP protein-encoding gene into the host cell chromosome is selected for by inclusion of 0.01-300 μM methotrexate in the

cell culture medium (as described in Ausubel et al., supra). This dominant selection can be accomplished in most cell types. Recombinant protein expression can be increased by DHFR-mediated amplification of the transfected gene. Methods for selecting cell lines bearing gene amplifications are described in Ausubel et al. (supra); such methods generally involve extended culture in medium containing gradually increasing levels of methotrexate. DHFR-containing expression vectors commonly used for this purpose include pCVSEII-DHFR and pAdD26SV(A) (described in Ausubel et al., supra). Any of the host cells described above or, preferably, a DHFR-deficient CHO cell line (e.g., CHO DHFR$^-$ cells, ATCC Accession No. CRL 9096) are among the host cells preferred for DHFR selection of a stably-transfected cell line or DHFR-mediated gene amplification.

Once the recombinant IAP protein is expressed, it is isolated, e.g., using affinity chromatography. In one example, an anti-IAP protein antibody (e.g., produced as described herein) may be attached to a column and used to isolate the IAP protein. Lysis and fractionation of IAP protein-harboring cells prior to affinity chromatography may be performed by standard methods (see, e.g., Ausubel et al., supra).

Once isolated, the recombinant protein can, if desired, be further purified, e.g., by high performance liquid chromatography (see, e.g., Fisher, Laboratory Techniques In Biochemistry And Molecular Biology, eds., Work and Burdon, Elsevier, 1980).

Polypeptides of the invention, particularly short IAP protein fragments, can also be produced by chemical synthesis (e.g., by the methods described in Solid Phase Peptide Synthesis, 2nd ed., 1984 The Pierce Chemical Co., Rockford, IL).

These general techniques of polypeptide expression and purification can also be used to produce and isolate useful IAP fragments or analogs (described herein).

VI. Anti-IAP Antibodies

To generate IAP-specific antibodies, a IAP coding sequence (i.e., amino acids 180-276) can be expressed as a C-terminal fusion with glutathione S-transferase (GST) (Smith et al., Gene 67:31-40, 1988). The fusion protein can be purified on glutathione-Sepharose beads, eluted with glutathione cleaved with thrombin (at the engineered cleavage site), and purified to the degree necessary for immunization of rabbits. Primary immunizations can be carried out with Freund's complete adjuvant and subsequent immunizations with Freund's incomplete adjuvant. Antibody titres are monitored by Western blot and immunoprecipitation analyses using the thrombin-cleaved IAP protein fragment of the GST-IAP fusion protein. Immune sera are affinity purified using CNBr-Sepharose-coupled IAP protein. Antiserum specificity is determined using a panel of unrelated GST proteins (including GSTp53, Rb, HPV-16 E6, and E6-AP) and GST-trypsin (which was generated by PCR using known sequences).

As an alternate or adjunct immunogen to GST fusion proteins, peptides corresponding to relatively unique hydrophilic regions of IAP may be generated and coupled to keyhole limpet hemocyanin (KLH) through an introduced C-terminal lysine. Antiserum to each of these peptides is similarly affinity purified on peptides conjugated to BSA, and specificity tested in ELISA and Western blots using peptide conjugates, and by Western blot and immunoprecipitation using IAP expressed as a GST fusion protein.

Alternatively, monoclonal antibodies may be prepared using the IAP proteins described above and standard hybridoma technology (see, e.g., Kohler et al., Nature 256:495, 1975; Kohler et al., Eur. J. Immunol. 6:511, 1976; Kohler et al., Eur. J. Immunol. 6:292, 1976; Hammerling et al., In Monoclonal Antibodies and T Cell Hybridomas, Elsevier, NY, 1981; Ausubel et al., supra). Once produced, monoclonal antibodies are also tested for specific IAP recognition by Western blot or immunoprecipitation analysis (by the methods described in Ausubel et al., supra). Antibodies which specifically recognize IAP are considered to be useful in the invention; such antibodies may be used, e.g., in an immunoassay to monitor the level of IAP produced by a mammal (for example, to determine the amount or subcellular location of IAP).

Preferably, antibodies of the invention are produced using fragments of the IAP protein which lie outside highly conserved regions and appear likely to be antigenic, by criteria such as those provided by the Peptidestructure program of the Genetics Computer Group Sequence Analysis Package (Program Manual for the GCG Package, Version 7, 1991) using the algorithm of Jameson and Wolf (CABIOS 4:181 1988)). Specifically these regions, which are found between BIR1 and BIR2 of all the IAP proteins, are in hiap1 from amino acid 99 to 170, hiap2 from amino acid 123 to 184, xiap from 116 to 133 and m-xiap from 116 to 133. In one specific example, such fragments are generated by standard techniques of PCR and cloned into the pGEX expression vector (Ausubel et al., supra). Fusion proteins are expressed in E. coli and purified using a glutathione agarose affinity matrix as described in Ausubel et al. (supra). To attempt to minimize the potential problems of low affinity or specificity of antisera, two or three such fusions are generated for each

protein, and each fusion is injected into at least two rabbits. Antisera are raised by injections in a series, preferably including at least three booster injections.

VII. Identification of Molecules that Modulate IAP Protein Expression

Isolation of the IAP cDNAs also facilitates the identification of molecules which increase or decrease IAP expression. According to one approach, candidate molecules are added at varying concentrations to the culture medium of cells expressing IAP mRNA. IAP expression is then measured, for example, by standard Northern blot analysis (Ausubel et al., supra) using a IAP cDNA (or cDNA fragment) as a hybridization probe (see also Table 5). The level of IAP expression in the presence of the candidate molecule is compared to the level measured for the same cells in the same culture medium but in the absence of the candidate molecule.

If desired, the effect of candidate modulators on expression may, in the alternative, be measured at the level of IAP protein production using the same general approach and standard immunological detection techniques, such as Western blotting or immunoprecipitation with a IAP-specific antibody (for example, the IAP antibody described herein).

Candidate modulators may be purified (or substantially purified) molecules or may be one component of a mixture of compounds (e.g., an extract or supernatant obtained from cells; Ausubel et al., supra). In a mixed compound assay, IAP expression is tested against progressively smaller subsets of the candidate compound pool (e.g., produced by standard purification techniques, e.g., HPLC or FPLC) until a single compound or minimal compound mixture is demonstrated to modulate IAP expression.

Alternatively, or in addition, candidate compounds may be screened for those which modulate IAP apoptosis inhibiting activity. In this approach, the degree of apoptosis in the presence of a candidate compound is compared to the degree of apoptosis in its absence, under equivalent conditions. Again, such a screen may begin with a pool of candidate compounds, from which one or more useful modulator compounds are isolated in a step-wise fashion. Apoptosis activity may be measured by any standard assay, for example, those described herein.

Another method for detecting compounds which modulate IAP polypeptide activity is to screen for compounds which physically interact with a given IAP polypeptide. Such compounds may be detected using adaptations of the interaction trap expression systems known in the art. Such systems detect protein interactions using a transcriptional activation assay and are generally described in Gyuris et al., Cell 75:791-803 (1993), and Field and Song, Nature 340:245-246, (1989), and are commercially available from Clonetech (Palo Alto, CA). In addition, PCT Publication WO 95/28497 (hereby incorporated by reference) describe a method for detecting proteins involved in apoptosis by virtue of their interaction with Bcl-2 using such an interaction trap assay. A similar method may be exploited to identify proteins and other compounds which interact with the IAP polypeptides.

Candidate IAP modulators include peptide as well as non-peptide molecules (e.g., peptide or non-peptide molecules found, e.g., in a cell extract, mammalian serum, or growth medium on which mammalian cells have been cultured).

A molecule which promotes an increase in IAP expression or IAP activity is considered particularly useful

in the invention; such a molecule may be used, for example, as a therapeutic to increase cellular levels of IAP and thereby exploit the effect of IAP polypeptides for the inhibition of apoptosis.

A molecule which decreases IAP activity (e.g., by decreasing gene expression or polypeptide activity) may be useful for decreasing cell proliferation. Such uses include treatment of neoplasms (see Table 3, below) or other cell proliferative diseases.

Modulators found to be effective at the level of IAP expression or activity may be confirmed as useful in animal models and, if successful, may be used as anti-cancer therapeutics for either the inhibition or the enhancement of apoptosis, as appropriate.

IX. IAP Therapy

Because expression levels of IAP genes correlates with the levels of apoptosis, the IAP gene also finds use in anti-apoptosis gene therapy. In particular, to sustain neuronal cells, lymphocytes (T-cells and B-cells), or cells exposed to ischemic injury, a functional IAP gene may be introduced into cells at the sites predicted to undergo undesirable apoptosis.

Retroviral vectors, adenoviral vectors, adeno-associated viral vectors, or other viral vectors with the appropriate tropism for cells likely to be involved in apoptosis (for example, epithelial cells) may be used as a gene transfer delivery system for a therapeutic IAP gene construct. Numerous vectors useful for this purpose are generally known (Miller, Human Gene Therapy 15-14, 1990; Friedman, Science 244:1275-1281, 1989; Eglitis and Anderson, BioTechniques 6:608-614, 1988; Tolstoshev and Anderson, Current Opinion in Biotechnology 1:55-61, 1990; Sharp, The

Lancet 337:1277-1278, 1991; Cornetta et al., Nucleic Acid Research and Molecular Biology 36:311-322, 1987; Anderson, Science 226:401-409, 1984; Moen, Blood Cells 17:407-416, 1991; and Miller and Rosman, Biotechniques 7:980-990, 1989; Le Gal La Salle et al., Science 259:988-990, 1993; and Johnson, Chest 107:77S-83S, 1995). Retroviral vectors are particularly well developed and have been used in clinical settings (Rosenberg et al., N. Engl. J. Med 323:370, 1990; Anderson et al., U.S. Pat. No. 5,399,346).

Non-viral approaches may also be employed for the introduction of therapeutic DNA into cells otherwise predicted to undergo apoptosis. For example, IAP may be introduced into a neuronal cell or a T-cell by the techniques of lipofection (Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413, 1987; Ono et al., Neuroscience Lett 117:259, 1990; Brigham et al., Am. J. Med. Sci. 298:278, 1989; Staubinger and Papahadjopoulos, Meth. Enz. 101:512, 1983); asialorosonucoid-polylysine conjugation (Wu and Wu, J. Biol. Chem. 263:14621, 1988; Wu et al., J. Biol. Chem. 264:16985, 1989); or, less preferably, microinjection under surgical conditions (Wolff et al., Science 247:1465, 1990).

For any of the above approaches, the therapeutic IAP DNA construct is preferably applied to the site of the predicted apoptosis event (for example, by injection), but may also be applied to tissue in the vicinity of the predicted apoptosis event or even to a blood vessel supplying the cells predicted to undergo apoptosis.

In the gene therapy constructs, IAP cDNA expression is directed from any suitable promoter (e.g., the human cytomegalovirus, simian virus 40, or metallothionein promoters), and its production is regulated by any desired mammalian regulatory element. For example, if desired, enhancers known to direct preferential gene expression in

neural cells or T-cells may be used to direct IAP expression. Such enhancers include, without limitation, those enhancers which are characterized as tissue or cell specific in their expression.

Alternatively, if a IAP genomic clone is utilized as a therapeutic construct (for example, following its isolation by hybridization with the IAP cDNA described above), IAP expression is regulated by its cognate regulatory sequences or, if desired, by regulatory sequences derived from a heterologous source, e.g., any of the promoters or regulatory elements described above.

Less preferably, IAP gene therapy is accomplished by direct administration of the IAP mRNA to a cell predicted to undergo apoptosis. This mRNA may be produced and isolated by any standard technique, but is most readily produced by in vitro transcription using a IAP cDNA under the control of a high efficiency promoter (e.g., the T7 promoter). Administration of IAP mRNA to malignant cells is carried out by any of the methods for direct nucleic acid administration described above.

Ideally, the production of IAP protein by any gene therapy approach described above results in a cellular level of IAP that is at least equivalent to the normal, cellular level of IAP in an unaffected individual. Treatment by any IAP-mediated gene therapy approach may be combined with more traditional therapies.

Another therapeutic approach included within the invention involves direct administration of recombinant IAP protein, either to the site of a predicted apoptosis event (for example, by injection) or systemically by any conventional recombinant protein administration technique. The actual dosage of IAP depends on a number of factors, including the size and health of the individual patient,

but, generally, between 0.1mg and 100mg inclusive are administered per day to an adult in any pharmaceutically-acceptable formulation.

X. Administration of IAP polypeptides, IAP genes, or modulators of IAP synthesis or function

A IAP protein, gene, or modulator may be administered with a pharmaceutically-acceptable diluent, carrier, or excipient, in unit dosage form. Conventional pharmaceutical practice may be employed to provide suitable formulations or compositions to administer IAP to patients suffering from or presymptomatic for a IAP-associated carcinoma. Any appropriate route of administration may be employed, for example, parenteral, intravenous, subcutaneous, intramuscular, intracranial, intraorbital, ophthalmic, intraventricular, intracapsular, intraspinal, intracisternal, intraperitoneal, intranasal, aerosol, or oral administration. Therapeutic formulations may be in the form of liquid solutions or suspensions; for oral administration, formulations may be in the form of tablets or capsules; and for intranasal formulations, in the form of powders, nasal drops, or aerosols.

Methods well known in the art for making formulations are found in, for example, "Remington's Pharmaceutical Sciences." Formulations for parenteral administration may, for example, contain excipients, sterile water, or saline, polyalkylene glycols such as polyethylene glycol, oils of vegetable origin, or hydrogenated napthalenes. Biocompatible, biodegradable lactide polymer, lactide/glycolide copolymer, or polyoxyethylene-polyoxypropylene copolymers may be used to control the release of the compounds. Other potentially useful parenteral delivery systems for IAP modulatory compounds

include ethylene-vinyl acetate copolymer particles, osmotic pumps, implantable infusion systems, and liposomes. Formulations for inhalation may contain excipients, for example, lactose, or may be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate and deoxycholate, or may be oily solutions for administration in the form of nasal drops, or as a gel.

If desired, treatment with a IAP protein, gene, or modulatory compound may be combined with more traditional therapies for the disease such as surgery, radiation, or chemotherapy for cancers; surgery, steroid therapy, and chemotherapy for autoimmune diseases; antiviral therapies for AIDS; and for example, TPA for ischemic injury.

XI. Detection of A Condition Involving Altered Apoptosis

IAP polypeptides and nucleic acid sequences find diagnostic use in the detection or monitoring of conditions involving aberrant levels of apoptosis. For example, decrease expression of IAP may be correlated with enhanced apoptosis in humans (see XII, below). Accordingly, a decrease or increase in the level of IAP production may provide an indication of a deleterious condition. Levels of IAP expression may be assayed by any standard technique. For example, its expression in a biological sample (e.g., a biopsy) may be monitored by standard Northern blot analysis or may be aided by PCR (see, e.g., Ausubel et al., supra; PCR Technology: Principles and Applications for DNA Amplification, ed., H.A. Ehrlich, Stockton Press, NY; and Yap and McGee, Nucl. Acids. Res. 19:4294, 1991).

Alternatively, a patient sample may be analyzed for one or more mutations in the IAP sequences using a mismatch detection approach. Generally, these techniques involve PCR amplification of nucleic acid from the patient sample,

followed by identification of the mutation (i.e., mismatch) by either altered hybridization, aberrant electrophoretic gel migration, binding or cleavage mediated by mismatch binding proteins, or direct nucleic acid sequencing. Any of these techniques may be used to facilitate mutant IAP detection, and each is well known in the art; examples of particular techniques are described, without limitation, in Orita et al., Proc. Natl. Acad. Sci. USA 86:2766-2770, (1989); and Sheffield et al., Proc. Natl. Acad. Sci. USA 86:232-236, (1989).

In yet another approach, immunoassays are used to detect or monitor IAP protein in a biological sample. IAP-specific polyclonal or monoclonal antibodies (produced as described above) may be used in any standard immunoassay format (e.g., ELISA, Western blot, or RIA assay) to measure IAP polypeptide levels; again comparison is to wild-type IAP levels, and a decrease in IAP production is indicative of a condition involving increased apoptosis. Examples of immunoassays are described, e.g., in Ausubel et al., supra. Immunohistochemical techniques may also be utilized for IAP detection. For example, a tissue sample may be obtained from a patient, and a section stained for the presence of IAP using an anti-IAP antibody and any standard detection system (e.g., one which includes a secondary antibody conjugated to horseradish peroxidase). General guidance regarding such techniques can be found in, e.g., Bancroft and Stevens (Theory and Practice of Histological Techniques, Churchill Livingstone, 1982) and Ausubel et al. (supra).

In one preferred example, a combined diagnostic method may be employed that begins with an evaluation of IAP protein production (for example, by immunological techniques or the protein truncation test (Hogerrorst, F.B.L., et al., Nature Genetics 10:208-212 (1995) and also includes a

nucleic acid-based detection technique designed to identify more subtle IAP mutations (for example, point mutations). As described above, a number of mismatch detection assays are available to those skilled in the art, and any preferred technique may be used (see above). By this approach, mutations in IAP may be detected that either result in loss of IAP expression or loss of IAP biological activity. In a variation of this combined diagnostic method, IAP biological activity is measured as protease activity using any appropriate protease assay system (for example, those described above).

Mismatch detection assays also provide the opportunity to diagnose a IAP-mediated predisposition to diseases of apoptosis. For example, a patient heterozygous for an IAP mutation may show no clinical symptoms and yet possess a higher than normal probability of developing one or more types of neurodegenerative, myelodysplastic or ischemic diseases. Given this diagnosis, a patient may take precautions to minimize their exposure to adverse environmental factors (for example, UV exposure or chemical mutagens) and to carefully monitor their medical condition (for example, through frequent physical examinations). This type of IAP diagnostic approach may also be used to detect IAP mutations in prenatal screens.

The IAP diagnostic assays described above may be carried out using any biological sample (for example, any biopsy sample or bodily fluid or tissue) in which IAP is normally expressed (for example, the inhibition of apoptosis). Identification of a mutant IAP gene may also be assayed using these sources for test samples. Alternatively, a IAP mutation, particularly as part of a diagnosis for predisposition to IAP-associated degenerative disease, may be tested using a DNA sample from any cell, for

example, by mismatch detection techniques; preferably, the DNA sample is subjected to PCR amplification prior to analysis.

To demonstrate the utility of IAP gene sequences as diagnostics and prognostics for cancer we probed the Clonetech (La Jolla) Human Cancer Cell Line Multiple Tissue Northern Blot (#7757-1). As Table 3 shows, all cancer lines tested showed increased IAP expression relative to samples from non-cancerous control cell lines. xiap expression was particularly high in HeLa (S-3), chronic myelogenous leukemia (K-562), colorectal adenocarcinoma (SW-480) and melanoma (G-361) lines. hiap1 expression was extremely high in Burkitt's lymphoma and was also elevated in colorectal adenocarcinoma. hiap2 expression was particularly high in chronic myelogenous leukemia (K-562) and colorectal adenocarcinoma (SW-480).

In addition, we note that we have mapped hiap1 and hiap2 to human chromosome 11q23. This is a known hotspot for cancer causing mutations.

TABLE 3

Northern Blot IAP RNA levels in Cancer Cells*

	xiap	hiap1	hiap2
Promylocytic Leukemia HL-60	+	+	+
Hela S-3	+	+	+
Chronic Myclogenous Leukemia K-562	+++	+	+++
Lymphoblastic Leukemia MDLT-4	+++	+	+
Burkitt's Lymphoma Raji	+	+(x10)	+
Colorectal Adenocarcinoma SW-480	+++	+++	+++
Lung Carcinoma A-549	+	+	+
Melanoma G-361	+++	+	+

*Levels are indicated by a (+) and are the approximate increase in RNA levels relative to Northern blots of RNA from non-cancerous control cell lines. A single plus indicates an estimated increase of at least 1-fold

XII. Treatment of HIV Infected Individuals

We have found that hiap1 and hiap 2 expression is decreased significantly in HIV infected human cells. This decrease precedes apoptosis. The result indicates that administration of HIAP1, HIAP2, genes encoding these proteins, or compounds which upregulate these genes can be used to prevent T-cell attrition in HIV infected patients. The following assay may also be used to screen for compounds

which alter hiap1 and hiap2 expression and which also prevent apoptosis.

The experiments were preformed as follows: Cultured mature lymphocyte CD-4$^+$ T-cell lines (H9 labelled "a"; CEM/CM-3 labelled "b"; 6T-CEM labelled "c"; and Jurkat labelled "d" in Figs. 13A and 13B) were examined for apoptosis (Fig. 13A) and hiap gene expression (Fig. 13B). Control conditions are labelled as lane 1 in Fig. 13A and Fig. 13B. Lane 2 shows the result 24 hours after PHA/PMH (phytohemagglutinin, phorbol ester) mitogen stimulation. Lane 3 shows the result 24 hours after HIV strain III_B infection. The "M" refers to standard DNA markers, the 123 bp ladder (Gibco-BRL) in Fig. 13B, and lambda HindIII ladder (Gibco-BRL) in Fig. A.

In Fig. 13A is a picture of ethidium bromide stained gel showing the presence of DNA ladders (as assayed by the test of Prigent et al., J. of Immun. Methods, 160:139-140, (1993), indicative of apoptosis. The sensitivity and degree of apoptosis of the four T-cell lines varies following mitogen stimulation and HIV infection.

For the experiment examining hiap gene expression, total RNA was prepared from the cultured cells and subject to a reverse transcriptase reaction using oligo-dT priming. The RT cDNA products were PCR amplified using specific primers (as shown in Table 5) for the detection of hiap2a, hiap2b and hiap 1. PCR conditions were routine (94°C melting for 1 minute, 55°C annealing for 2 minutes and 72°C extension for 1.5 minutes for 35 cycles) using a Perkin-Elmer 480 thermocycler. The Fig. 13B shows a picture of the RT-PCR products run on a 1% agarose gel stained with ethidium bromide. Absence of hiap2 transcripts is noted in all four cell lines 24 hours after HIV infection. In three of four cell lines (all except H9), the hiap1 gene is also

dramatically down-regulated after HIV infection. PHA/PMA mitogen stimulation also appears to decrease hiap gene expression, particularly for hiap2 and to a lesser extent, for hiap1.

The data from these experiments is summarized in the accompanying Table 5. The ß-action gene expression was consistent in all cell lines tested, indicating that a flow in the RT-PCR assay does not account for the decreases in hiap gene expression.

Table 4

Oligonucleotide primers for the specific RT-PCR amplification of unique IAP genes.

IAP Gene	Forward Primer (nucleotide position*)	Reverse Primer (nucleotide position*)	Size of Product (bp)
h-xiap	p2415 (876-896)	p2449 (1291-1311)	435
m-xiap	p2566 (458-478)	p2490 (994-1013)	555
h-hiap1	p2465 (827-847)	p2464 (1008-1038)	211
m-hiap1	p2687 (747-767)	p2684 (1177-1197)	450
hiap2	p2595 (1562-1585)	p2578 (2339-2363)	801[a] 618[b]
m-hiap2	p2693 (1751-1772)	p2734 (2078-2100)	349

* Nucleotide position as determined from Figs. 1-4 for each IAP gene

[a] PCR product size of hiap2a

[b] PCR product size of hiap2b

Table 5

Apoptosis and hiap gene expression in cultured T-cells following mitogen stimulation or HIV infection.

Cell Line	Condition	Apoptosis	hiap1	hiap2
H9	not stimulated	-	+	+/-
	PHA/PMA stimulated	+++	+	+/-
	HIV infected	++	+	-
CEM/CM-3	not stimulated	-	+	+/-
	PHA/PMA stimulated	+/-	+	-
	HIV infected	+/-	-	-
6T-CEM	not stimulated	-	+	+
	PHA/PMA stimulated	+/-	-	-
	HIV infected	+	-	-
Jurkat	not stimulated	-	+	++
	PHA/PMA stimulated	+	+	+
	HIV infected	+/-	-	-

XIII. Preventive Anti-Apoptotic Therapy

In a patient diagnosed to be heterozygous for an IAP mutation or to be susceptible to IAP mutations (even if those mutations do not yet result in alteration or loss of IAP biological activity), or a patent diagnosed as HIV positive, any of the above therapies may be administered before the occurrence of the disease phenotype. For example, the therapies may be provided to a patient who is HIV positive but does not yet show a diminished T-cell count or other signs of full-blown AIDS. In particular, compounds shown to increase IAP expression or IAP biological activity may be administered by any standard dosage and route of administration (see above). Alternatively, gene therapy using an IAP expression construct may be undertaken to

reverse or prevent the cell defect prior to the development of the degenerative disease.

The methods of the instant invention may be used to reduce or diagnose the disorders described herein in any mammal, for example, humans, domestic pets, or livestock. Where a non-human mammal is treated or diagnosed, the IAP polypeptide, nucleic acid, or antibody employed is preferably specific for that species.

Other Embodiments

In other embodiments, the invention includes any protein which is substantially identical to a mammalian IAP polypeptides (Figs. 1-6; SEQ ID NO:1-42); such homologs include other substantially pure naturally-occurring mammalian IAP proteins as well as allelic variants; natural mutants; induced mutants; DNA sequences which encode proteins and also hybridize to the IAP DNA sequences of Figs. 1-6 (SEQ ID NOS:1-42) under high stringency conditions or, less preferably, under low stringency conditions (e.g., washing at 2X SSC at 40°C with a probe length of at least 40 nucleotides); and proteins specifically bound by antisera directed to a IAP polypeptide. The term also includes chimeric polypeptides that include a IAP portion.

The invention further includes analogs of any naturally-occurring IAP polypeptide. Analogs can differ from the naturally-occurring IAP protein by amino acid sequence differences, by post-translational modifications, or by both. Analogs of the invention will generally exhibit at least 85%, more preferably 90%, and most preferably 95% or even 99% identity with all or part of a naturally-occurring IAP amino acid sequence. The length of sequence comparison is at least 15 amino acid residues, preferably at least 25 amino acid residues, and more preferably more than

35 amino acid residues. Modifications include *in vivo* and *in vitro* chemical derivatization of polypeptides, e.g., acetylation, carboxylation, phosphorylation, or glycosylation; such modifications may occur during polypeptide synthesis or processing or following treatment with isolated modifying enzymes. Analogs can also differ from the naturally-occurring IAP polypeptide by alterations in primary sequence. These include genetic variants, both natural and induced (for example, resulting from random mutagenesis by irradiation or exposure to ethanemethylsulfate or by site-specific mutagenesis as described in Sambrook, Fritsch and Maniatis, Molecular Cloning: A Laboratory Manual (2d ed.), CSH Press, 1989, or Ausubel et al., *supra*). Also included are cyclized peptides, molecules, and analogs which contain residues other than L-amino acids, e.g., D-amino acids or non-naturally occurring or synthetic amino acids, e.g., ß or γ amino acids.

In addition to full-length polypeptides, the invention also includes IAP polypeptide fragments. As used herein, the term "fragment," means at least 20 contiguous amino acids, preferably at least 30 contiguous amino acids, more preferably at least 50 contiguous amino acids, and most preferably at least 60 to 80 or more contiguous amino acids. Fragments of IAP polypeptides can be generated by methods known to those skilled in the art or may result from normal protein processing (e.g., removal of amino acids from the nascent polypeptide that are not required for biological activity or removal of amino acids by alternative mRNA splicing or alternative protein processing events).

Preferable fragments or analogs according to the invention are those which facilitate specific detection of a IAP nucleic acid or amino acid sequence in a sample to be

diagnosed. Particularly useful IAP fragments for this purpose include, without limitation, the amino acid fragments shown in Table 2.

All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each independent publication or patent application was specifically and individually indicated to be incorporated by reference.

What is claimed is:

Claims

1. Substantially pure nucleic acid encoding an IAP polypeptide.

2. The nucleic acid of claim 1, wherein said polypeptide comprises a ring zinc finger domain and at least one BIR domain.

3. The nucleic acid of claim 2, wherein said polypeptide has at least two BIR domains.

4. The nucleic acid of claim 3, wherein said polypeptide has at least three BIR domains.

5. The nucleic acid of claim 1, wherein said DNA contains the xiap gene.

6. The nucleic acid of claim 1, wherein said DNA contains the hiap2 gene.

7. The nucleic acid of claim 1, wherein said DNA contains the hiap1 gene.

8. The nucleic acid of claim 1, wherein said nucleic acid is genomic DNA.

9. The nucleic acid of claim 1, wherein said nucleic acid is cDNA.

10. The nucleic acid of claim 1, wherein said nucleic acid is mammalian DNA.

11. The nucleic acid of claim 10, wherein said mammalian DNA is human DNA.

12. The nucleic acid of claim 10, wherein said mammalian DNA is murine DNA.

13. Substantially pure DNA having the sequence of Fig. 5, or degenerate variants thereof, and encoding the amino acid sequence of Fig. 5.

14. Substantially pure DNA having the sequence of Fig. 6, or degenerate variants thereof, and encoding the amino acid sequence of Fig. 6.

15. Substantially pure DNA having about 50% or greater nucleotide sequence identity to the DNA sequence of Fig. 5.

16. Substantially pure DNA having about 50% or greater nucleotide sequence identity to the DNA sequence of Fig. 6.

17. A purified DNA sequence substantially identical to the DNA sequence shown in Fig. 5.

18. A purified DNA sequence substantially identical to the DNA sequence shown in Fig. 6.

19. A substantially pure mammalian IAP polypeptide.

20. The polypeptide of claim 19, wherein said polypeptide is the murine HIAP1 polypeptide.

21. The polypeptide of claim 19, wherein said polypeptide is the murine HIAP2 polypeptide.

22. The polypeptide of claim 19, comprising an amino acid sequence substantially identical to an amino acid sequence shown in Fig. 5.

23. The polypeptide of claim 19, comprising an amino acid sequence substantially identical to an amino acid sequence shown in Fig. 6.

24. A therapeutic composition comprising as an active ingredient an IAP polypeptide according to claim 19, said active ingredient being formulated in a physiologically acceptable carrier.

25. A method of inhibiting apoptosis in a mammal, said method comprising:

providing a cell of said mammal with a transgene encoding an IAP polypeptide, said DNA positioned for expression in said cell.

26. The method of claim 25 wherein said polypeptide is murine HIAP1.

27. The method of claim 25 wherein said polypeptide is murine HIAP2.

28. A method of detecting an IAP gene in an animal cell, said method comprising:

contacting the DNA of claim 13 or a portion thereof greater than about 18 nucleic acids in length with a preparation of genomic DNA from said animal cell under

hybridization conditions providing detection of DNA sequences having about 50% or greater nucleotide sequence identity to the sequence of Fig. 5.

29. A method of detecting an IAP gene in an animal cell, said method comprising:

contacting the DNA of claim 14 or a portion thereof greater than about 18 nucleic acids in length with a preparation of genomic DNA from said animal cell under hybridization conditions providing detection of DNA sequences having about 50% or greater nucleotide sequence identity to the sequence of Fig. 6.

30. A method of producing an IAP polypeptide comprising:

providing a cell transformed with DNA encoding an IAP polypeptide positioned for expression in said cell; culturing said transformed cell under conditions for expressing said DNA; and

isolating said IAP polypeptide.

31. The method of claim 30, wherein said IAP polypeptide is murine HIAP1.

32. The method of claim 30, wherein said IAP polypeptide is murine HIAP2.

33. A method of identifying a compound which modulates apoptosis, said method comprising (a) providing a cell expressing an IAP polypeptide; and (b) contracting said cell with a candidate compound and monitoring the expression of an IAP gene, an alteration in the level of expression of

said gene indicating the presence of a compound which modulates apoptosis.

34. The method of claim 33, wherein said IAP gene is murine HIAP1.

35. The method of claim 33, wherein said IAP gene is murine HIAP2.

36. A method for detecting a protein that interacts with an IAP polypeptide comprising the steps of:

a. contacting under suitable conditions an IAP protein with a compound suspected to be a modulator of apoptosis; and

b. detecting the interaction of said compound with said IAP polypeptide, wherein said interaction indicates that said compound is involved in the modulation of apoptosis.

37. The method of claim 36, wherein said IAP polypeptide is HIAP1.

38. The method of claim 36, wherein said IAP polypeptide is HIAP2.

39. The method of claim 36, wherein said IAP polypeptide is XIAP.

40. The method of claim 36, wherein said interaction is detected by measuring the transcriptional activity of a reporter gene.

41. The method of claim 36, wherein said interaction occurs in a yeast cell.

42. The method of claim 36, wherein said compound is a polypeptide.

43. The method of claim 42, wherein said polypeptide is expressed from a recombinant nucleic acid.

44. A method of diagnosing an increased liklihood of a cell proliferative disease in a patient, said method comprising detecting the level of IAP gene expression in said patient.

45. A method of diagnosing an increased liklihood of a cell proliferative disease in a patient, said method comprising detecting the level of IAP polypeptide activity in said patient.

46. A transgenic rodent having a knockout mutation in an IAP gene.

47. A transgenic rodent, said rodent having additional copies of IAP nucleic acids added to its genome.

MAMMALIAN IAP GENE FAMILY, PRIMERS, PROBES, AND DETECTION METHODS

ABSTRACT OF THE DISCLOSURE

Disclosed is substantially pure DNA encoding mammalian IAP polypeptides; substantially pure polypeptides; and methods of using such DNA to express the IAP polypeptides in cells and animals to inhibit apoptosis. Also disclosed are conserved regions characteristic of the IAP family and primers and probes for the identification and isolation of additional IAP genes. In addition, methods for treating diseases and disorders involving apoptosis are provided.

159472.B11

HUMAN xiap

SEQ ID NO:3

```
        1 gaaaaggtggacaagtcctaatttcaagagaagatgacttttaacagttttgaaggatct
          ---------+---------+---------+---------+---------+---------+ 60
```

SEQ ID NO:4 a

```
                                          M  T  F  N  S  F  E  G  S  -
       61 aaaacttgtgtacctgcagacatcaataaggaagaagaatttgtagaagagtttaataga
          ---------+---------+---------+---------+---------+---------+ 120
a            K  T  C  V  P  A  D  I  N  K  E  E  E  F  V  E  E  F  N  R  -
      121 ttaaaaacttttgctaattttccaagtggtagtcctgtttcagcatcaacactggcacga
          ---------+---------+---------+---------+---------+---------+ 180
a            L  K  T  F  A  N  F  P  S  G  S  P  V  S  A  S  T  L  A  R  -
      181 gcagggtttctttatactggtgaaggagataccgtgcggtgctttagttgtcatgcagct
          ---------+---------+---------+---------+---------+---------+ 240
a            A  G  F  L  Y  T  G  E  G  D  T  V  R  C  F  S  C  H  A  A  -
      241 gtagatagatggcaatatggagactcagcagttggaagacacaggaaagtatccccaaat
          ---------+---------+---------+---------+---------+---------+ 300
a            V  D  R  W  Q  Y  G  D  S  A  V  G  R  H  R  K  V  S  P  N  -
      301 tgcagatttatcaacggcttttatcttgaaaatagtgccacgcagtctacaaattctggt
          ---------+---------+---------+---------+---------+---------+ 360
a            C  R  F  I  N  G  F  Y  L  E  N  S  A  T  Q  S  T  N  S  G  -
```

FIG. 1 (PAGE 1 OF 7)

HUMAN xiap

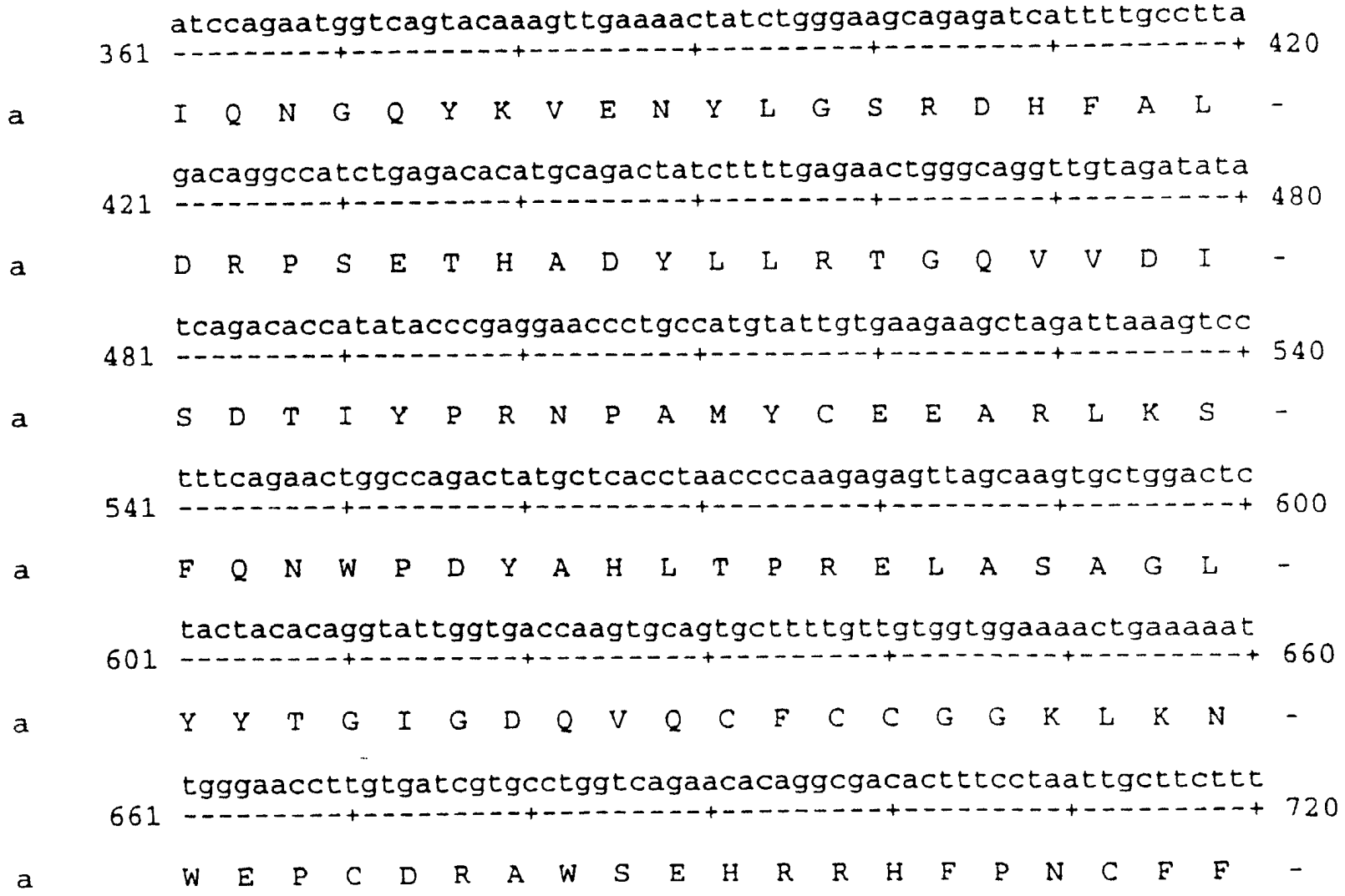

FIG. 1 (PAGE 2 OF 7)

HUMAN xiap

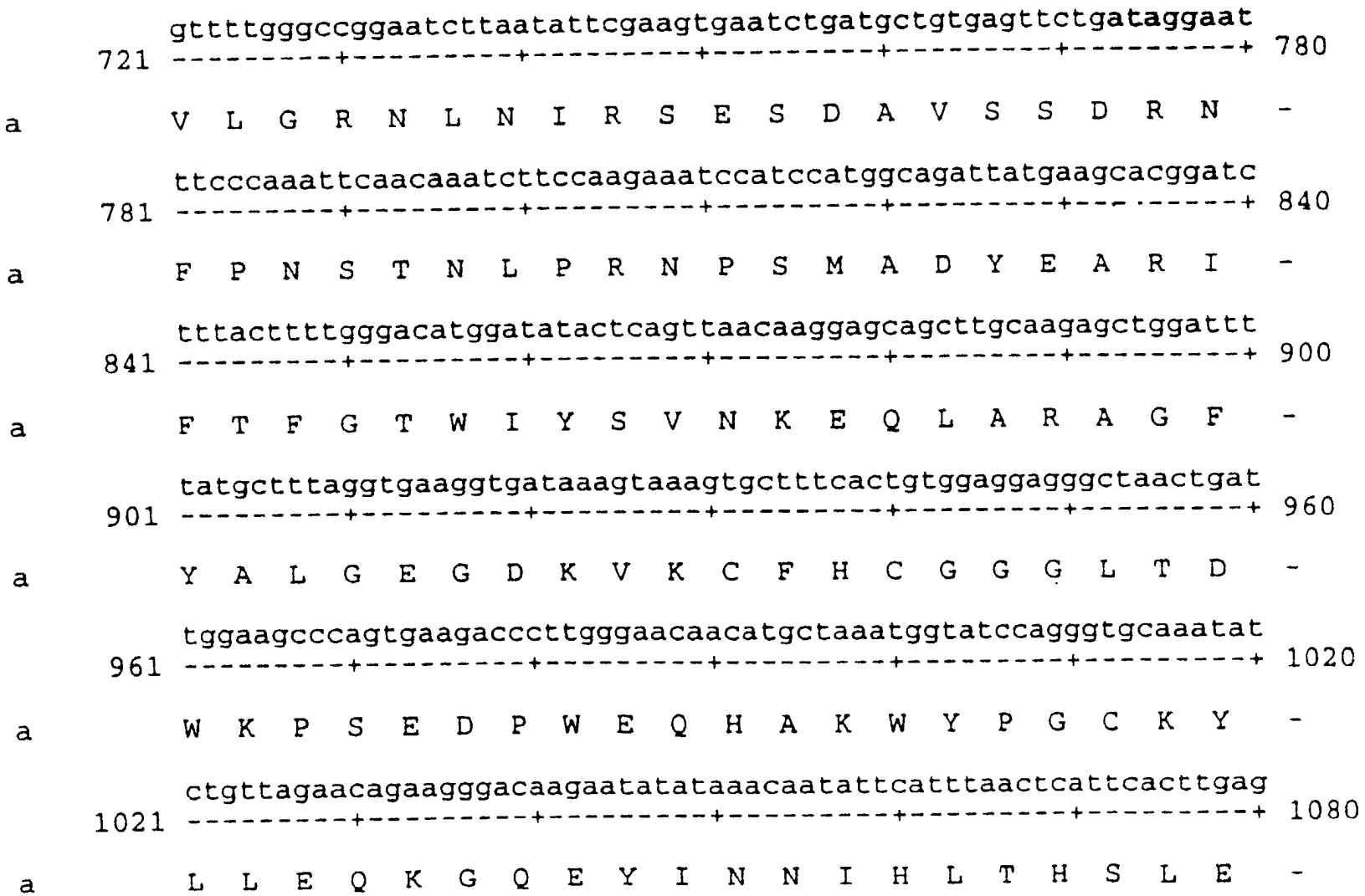

FIG. 1 (PAGE 3 OF 7)

HUMAN xiap

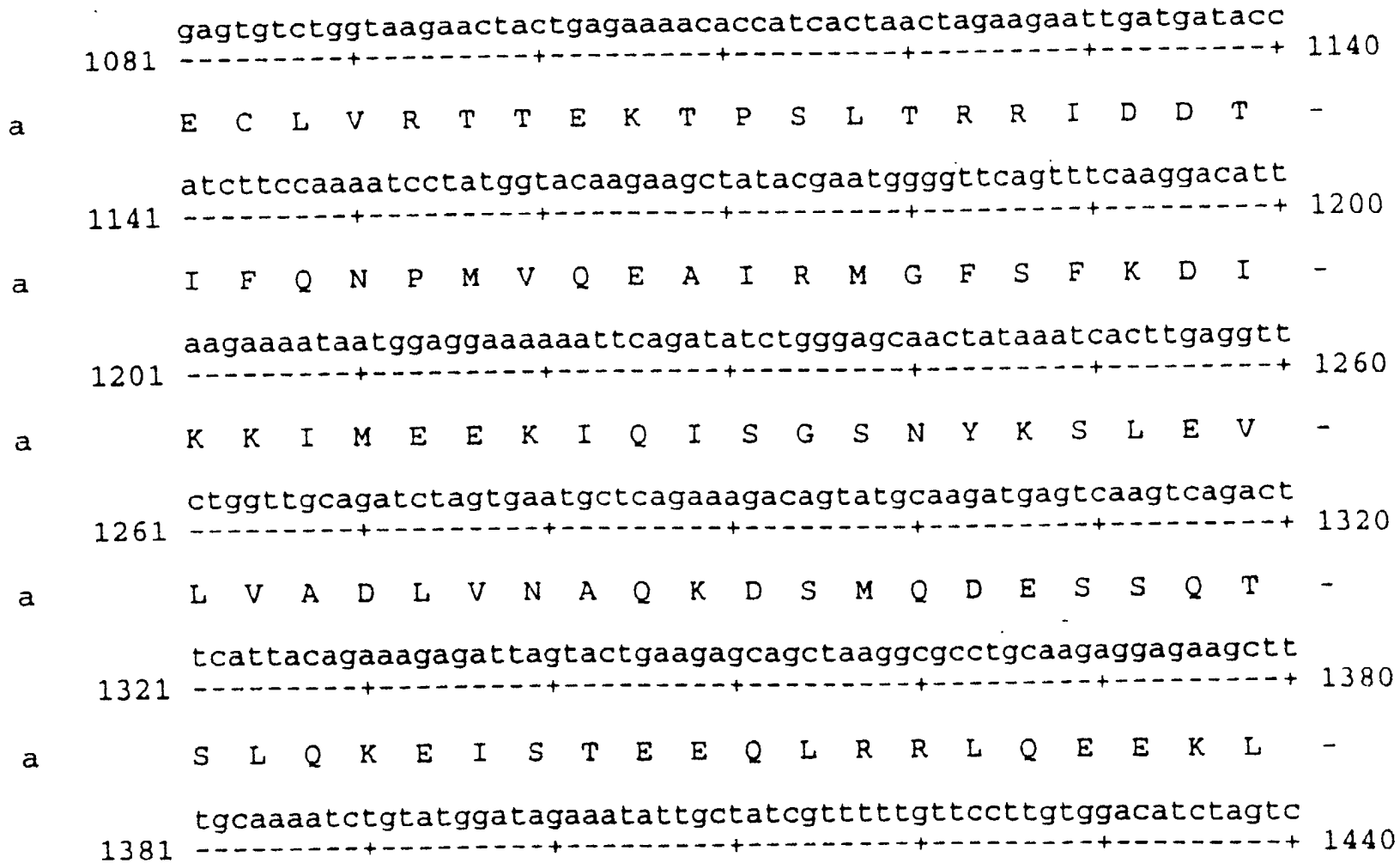

FIG. 1 (PAGE 4 OF 7)

HUMAN xiap

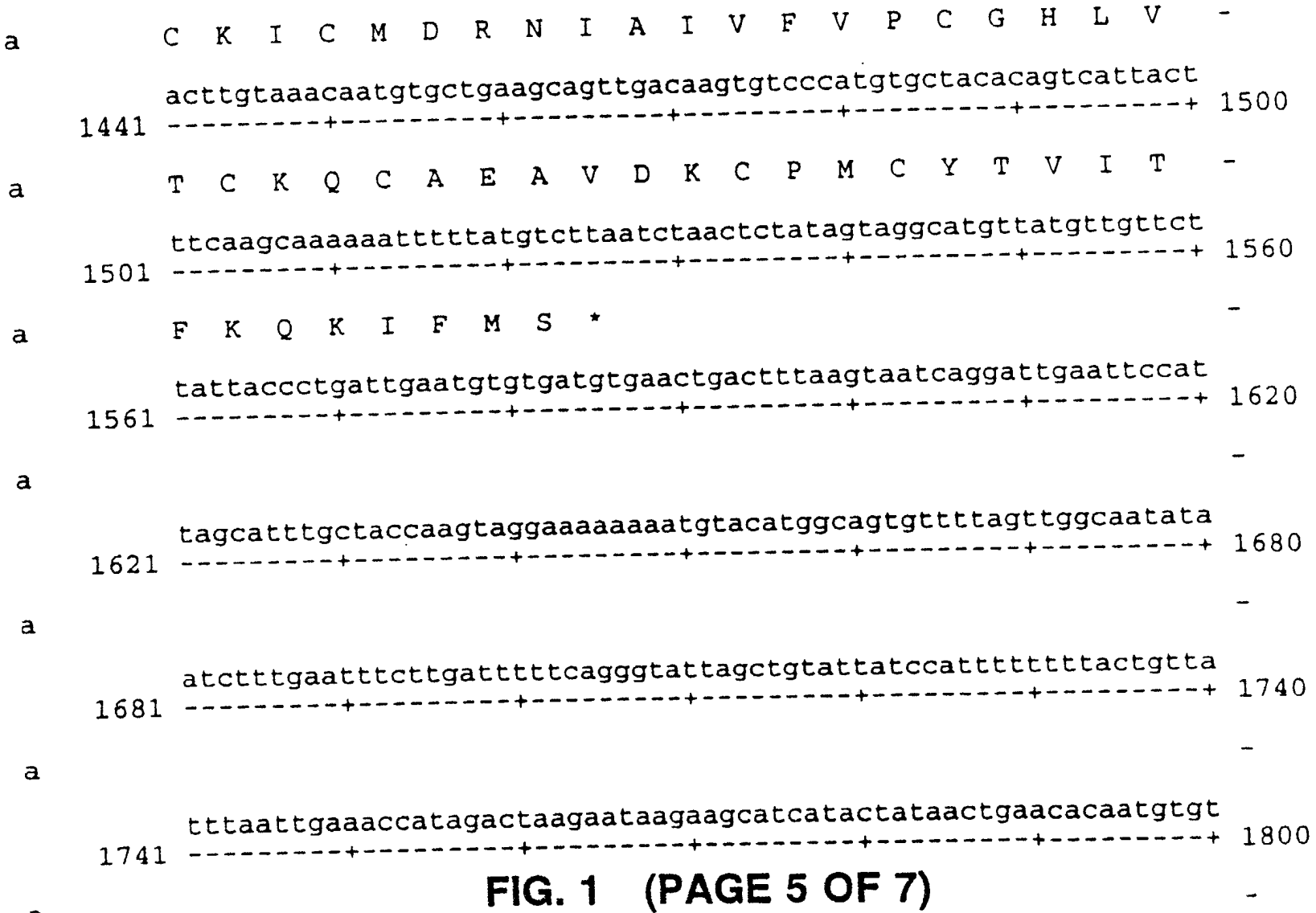

FIG. 1 (PAGE 5 OF 7)

HUMAN xiap

FIG. 1 (PAGE 6 OF 7)

HUMAN xiap

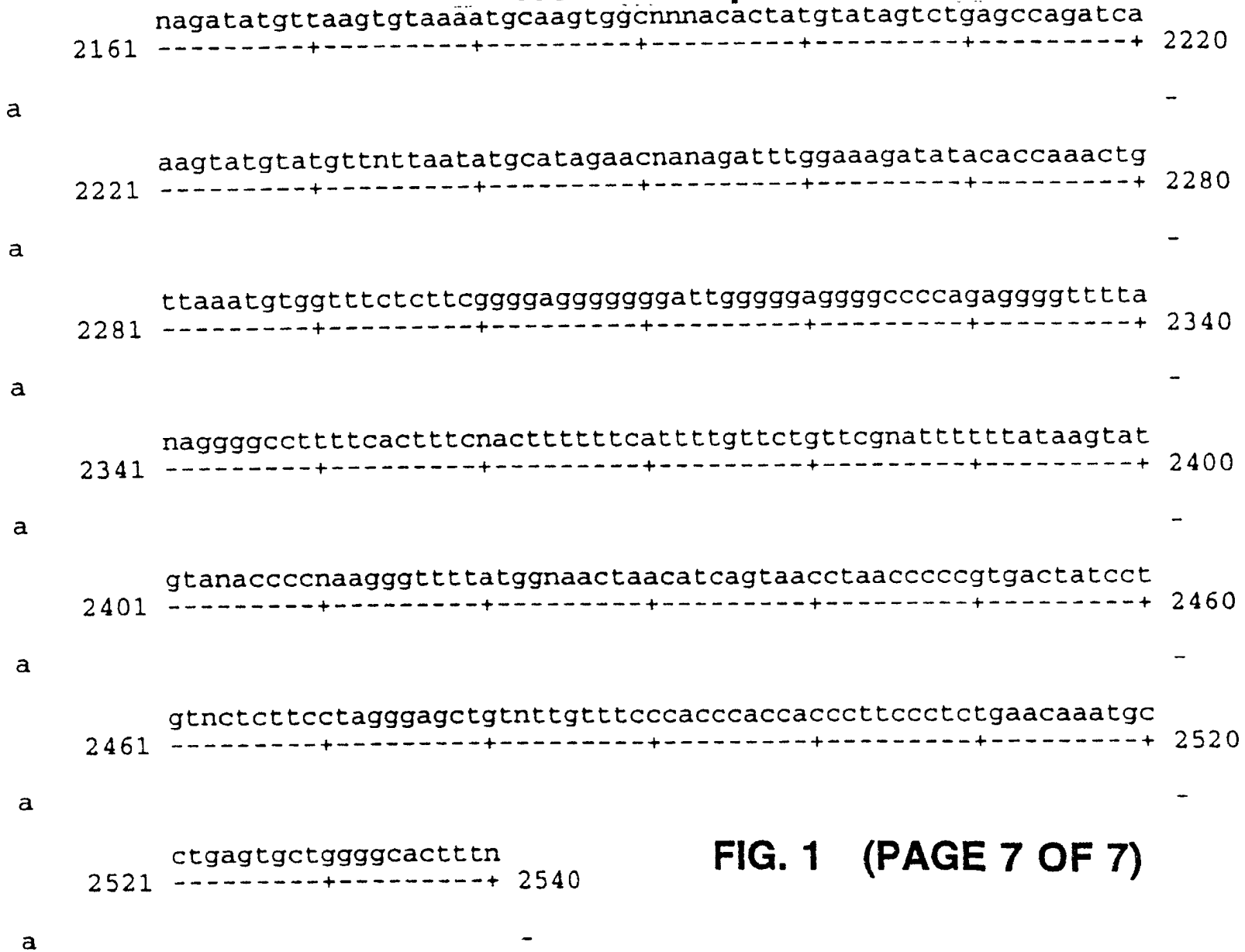

FIG. 1 (PAGE 7 OF 7)

HUMAN hiap-1

SEQ ID NO:5

```
      TCCTTGAGATGTATCAGTATAGGATTTAGGATCTCCATGTTGGAACTCTAAATGCATAGA
    1 ---------+---------+---------+---------+---------+---------+ 60
C                                                                   -

      AATGGAAATAATGGAAATTTTTCATTTTGGCTTTTCAGCCTAGTATTAAAACTGATAAAA
   61 ---------+---------+---------+---------+---------+---------+ 120
C                                                                   -

      GCAAAGCCATGCACAAAACTACCTCCCTAGAGAAAGGCTAGTCCCTTTTCTTCCCCATTC
  121 ---------+---------+---------+---------+---------+---------+ 180
C                                                                   -

      ATTTCATTATGAACATAGTAGAAAACAGCATATTCTTATCAAATTTGATGAAAAGCGCCA
  181 ---------+---------+---------+---------+---------+---------+ 240
SEQ ID NO:6 C       M  N  I  V  E  N  S  I  F  L  S  N  L  M  K  S  A  N -

      ACACGTTTGAACTGAAATACGACTTGTCATGTGAACTGTACCGAATGTCTACGTATTCCA
  241 ---------+---------+---------+---------+---------+---------+ 300
C       T  F  E  L  K  Y  D  L  S  C  E  L  Y  R  M  S  T  Y  S  T -

      CTTTTCCTGCTGGGGTTCCTGTCTCAGAAAGGAGTCTTGCTCGTGCTGGTTTCTATTACA
  301 ---------+---------+---------+---------+---------+---------+ 360
C       F  P  A  G  V  P  V  S  E  R  S  L  A  R  A  G  F  Y  Y  T -
```

FIG. 2 (PAGE 1 OF 8)

HUMAN hiap-1

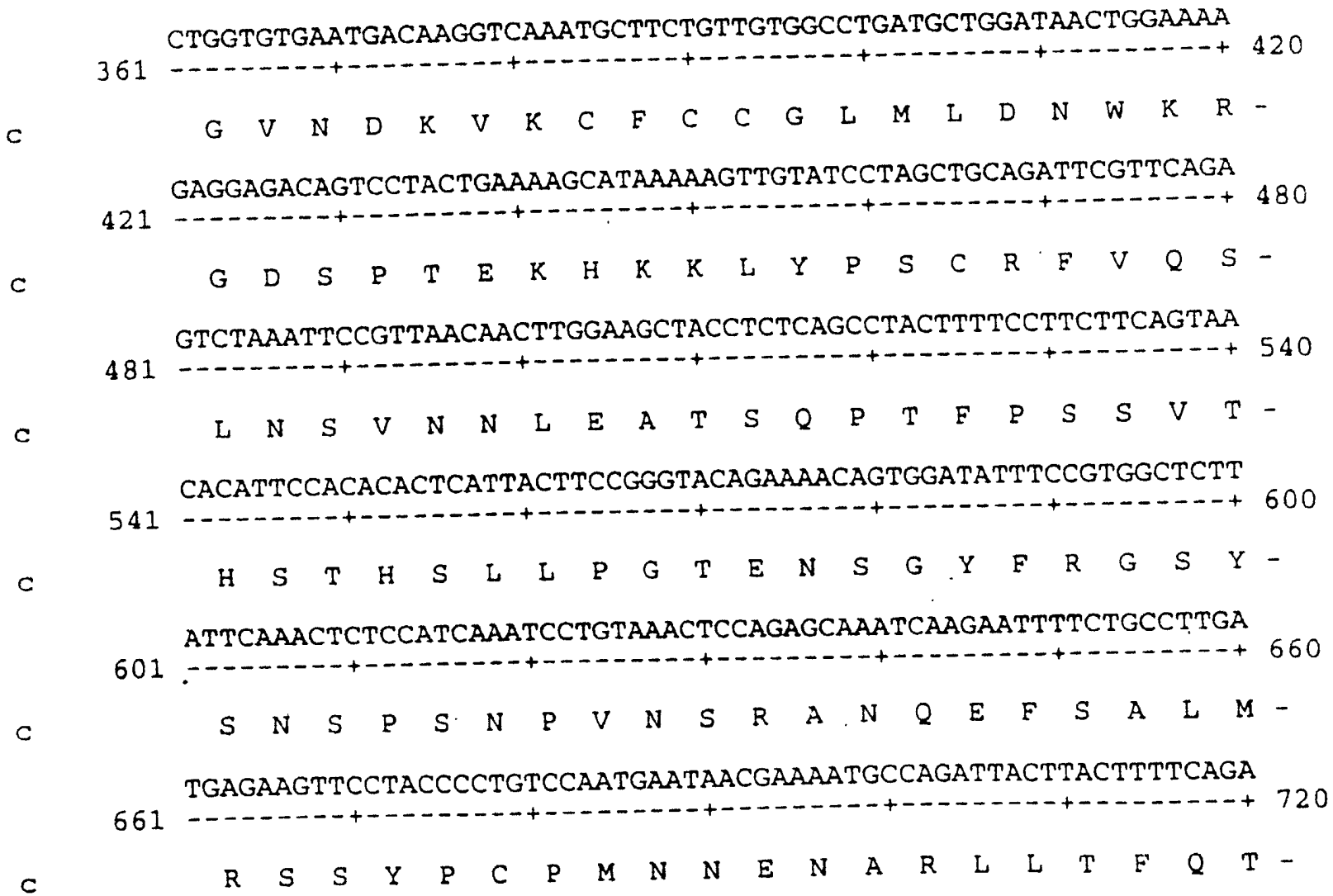

FIG. 2 (PAGE 2 OF 8)

HUMAN hiap-1

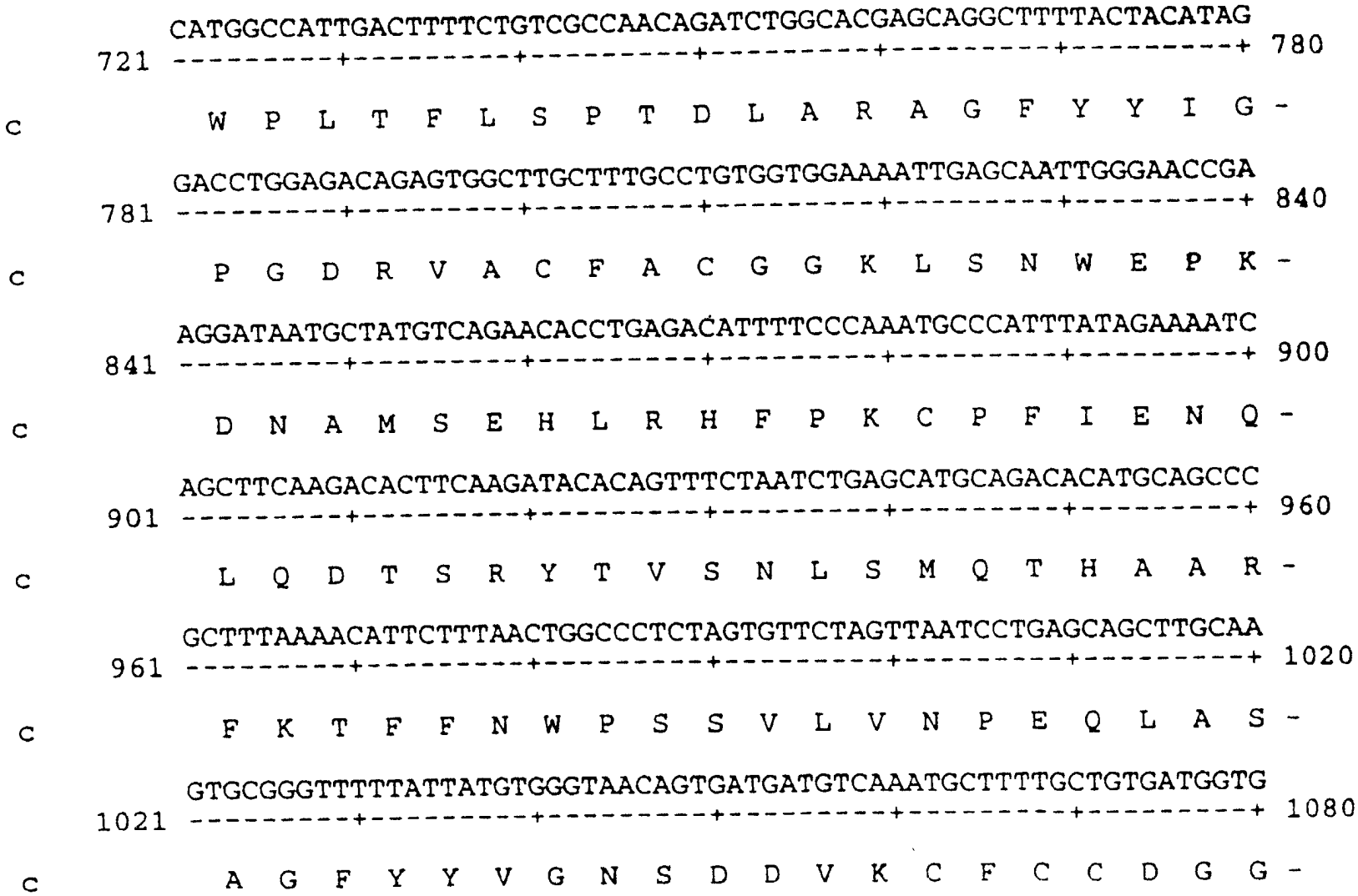

FIG. 2 (PAGE 3 OF 8)

HUMAN hiap-1

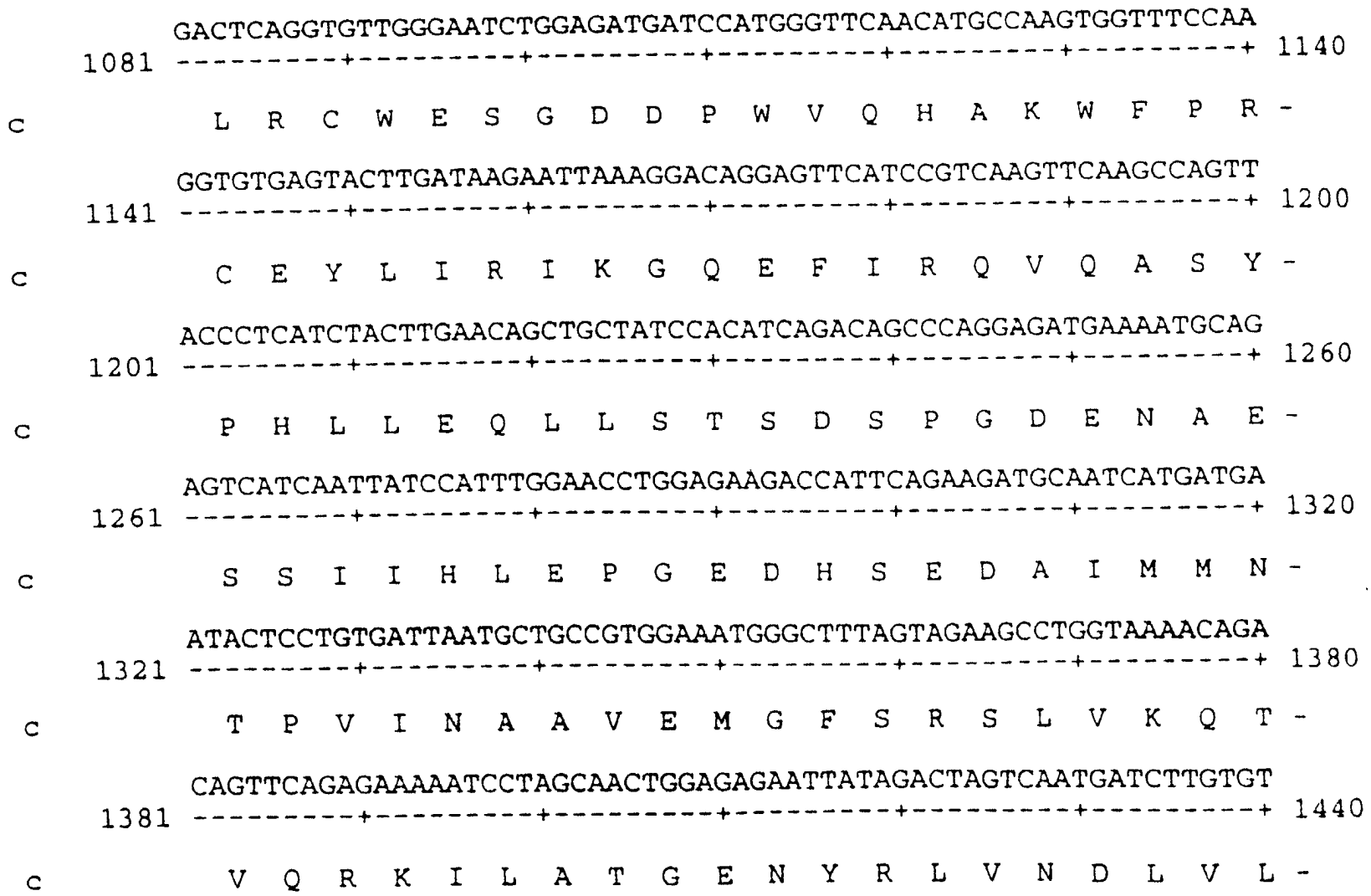

FIG. 2 (PAGE 4 OF 8)

HUMAN hiap-1

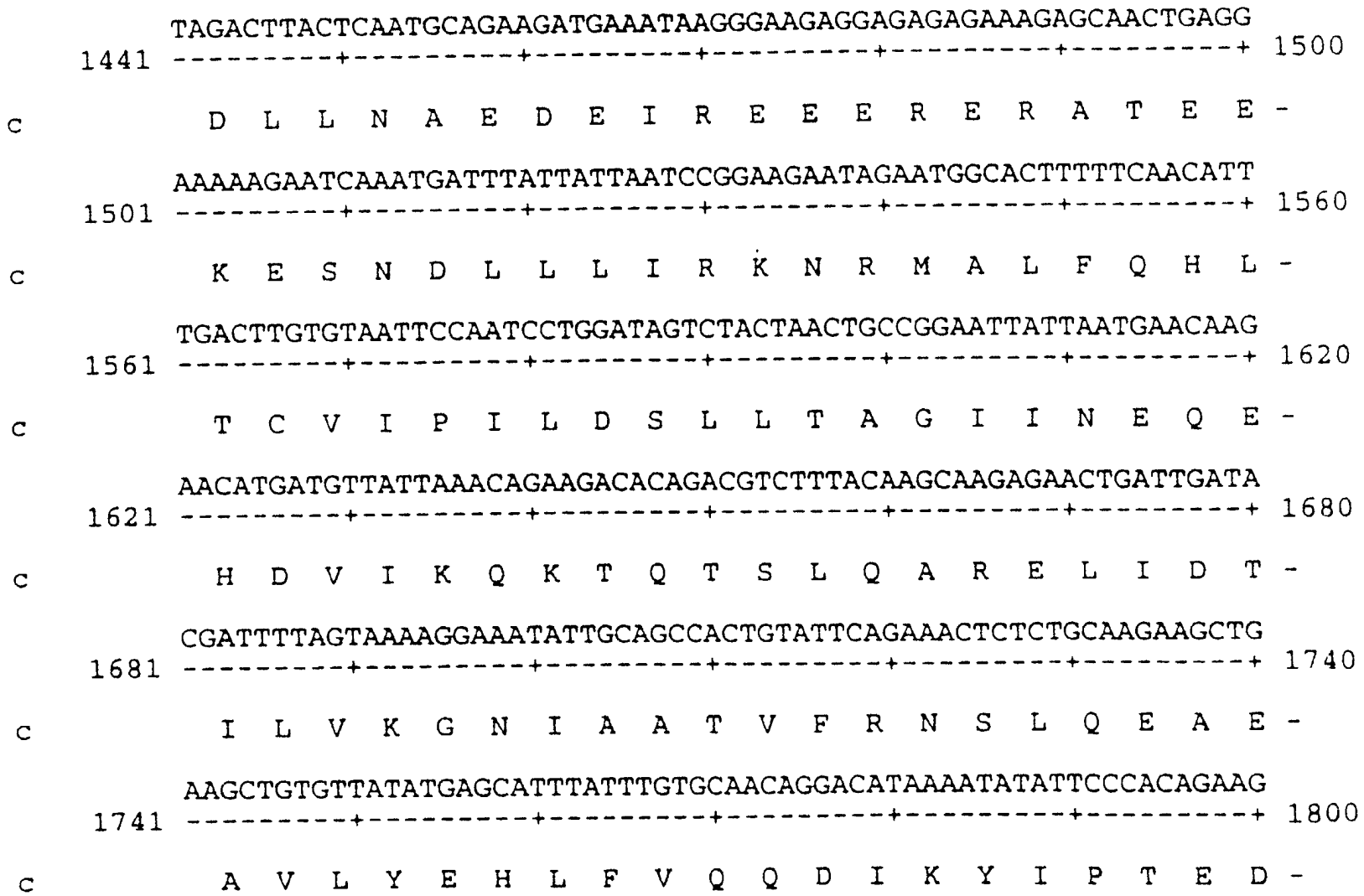

FIG. 2 (PAGE 5 OF 8)

HUMAN hiap-1

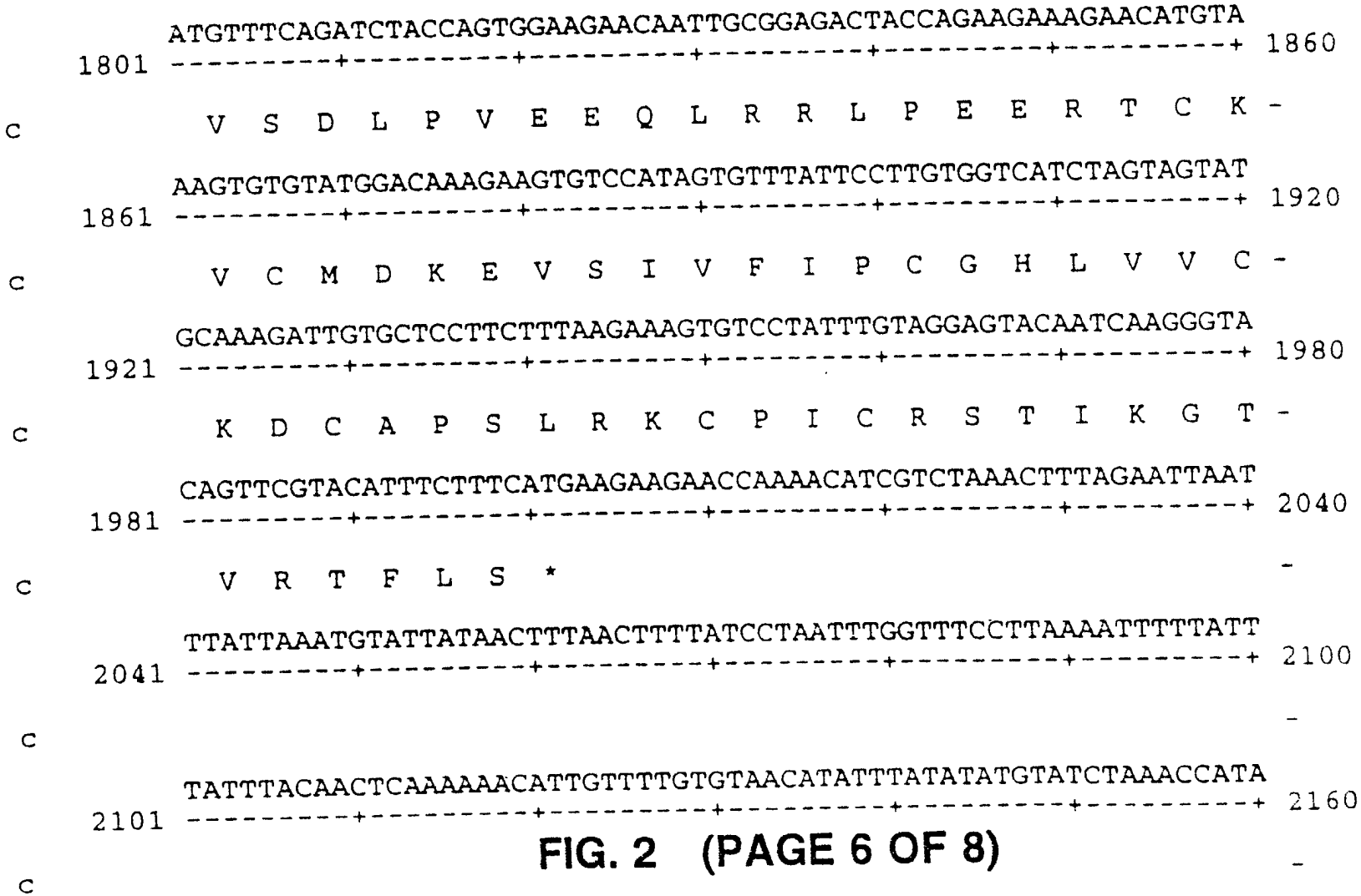

FIG. 2 (PAGE 6 OF 8)

HUMAN hiap-1

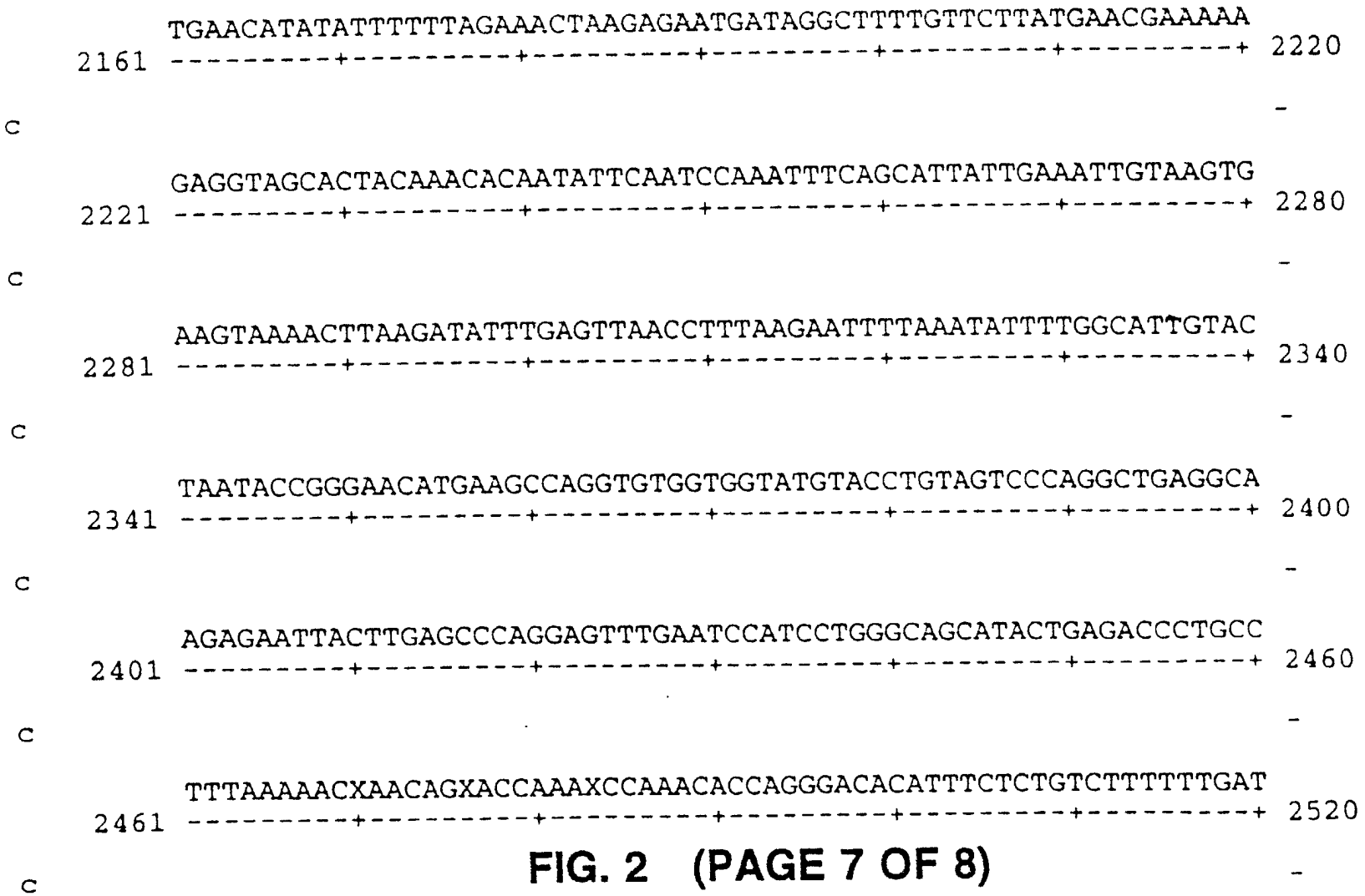

FIG. 2 (PAGE 7 OF 8)

HUMAN hiap-1

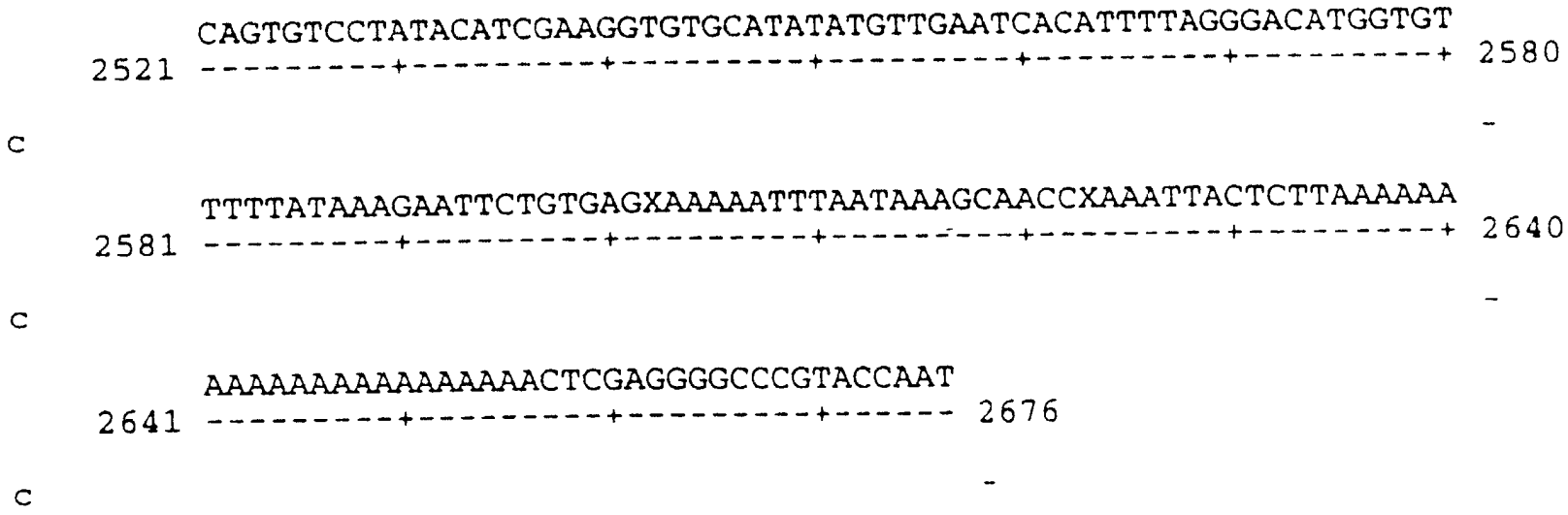

FIG. 2 (PAGE 8 OF 8)

HUMAN hiap-2

SEQ ID NO:7

```
          TTAGGTTACCTGAAAGAGTTACTACAACCCCAAAGAGTTGTGTTCTAAGTAGTATCTTGG
        1 ---------+---------+---------+---------+---------+---------+ 60

     a                                                                  -

          TAATTCAGAGAGATACTCATCCTACCTGAATATAAACTGAGATAAATCCAGTAAAGAAAG
       61 ---------+---------+---------+---------+---------+---------+ 120

     a                                                                  -

          TGTAGTAAATTCTACATAAGAGTCTATCATTGATTTCTTTTTGTGGTGGAAATCTTAGTT
      121 ---------+---------+---------+---------+---------+---------+ 180

     a                                                                  -

          CATGTGAAGAAATTTCATGTGAATGTTTTAGCTATCAAACAGTACTGTCACCTACTCATG
      181 ---------+---------+---------+---------+---------+---------+ 240

     a                                                            M     -

          CACAAAACTGCCTCCCAAAGACTTTTCCCAGGTCCCTCGTATCAAAACATTAAGAGTATA
      241 ---------+---------+---------+---------+---------+---------+ 300

SEQ ID NO:8 a     H  K  T  A  S  Q  R  L  F  P  G  P  S  Y  Q  N  I  K  S  I     -

          ATGGAAGATAGCACGATCTTGTCAGATTGGACAAACAGCAACAAACAAAAAATGAAGTAT
      301 ---------+---------+---------+---------+---------+---------+ 360

     a     M  E  D  S  T  I  L  S  D  W  T  N  S  N  K  Q  K  M  K  Y     -
```

FIG. 3 (PAGE 1 OF 7)

HUMAN hiap-2

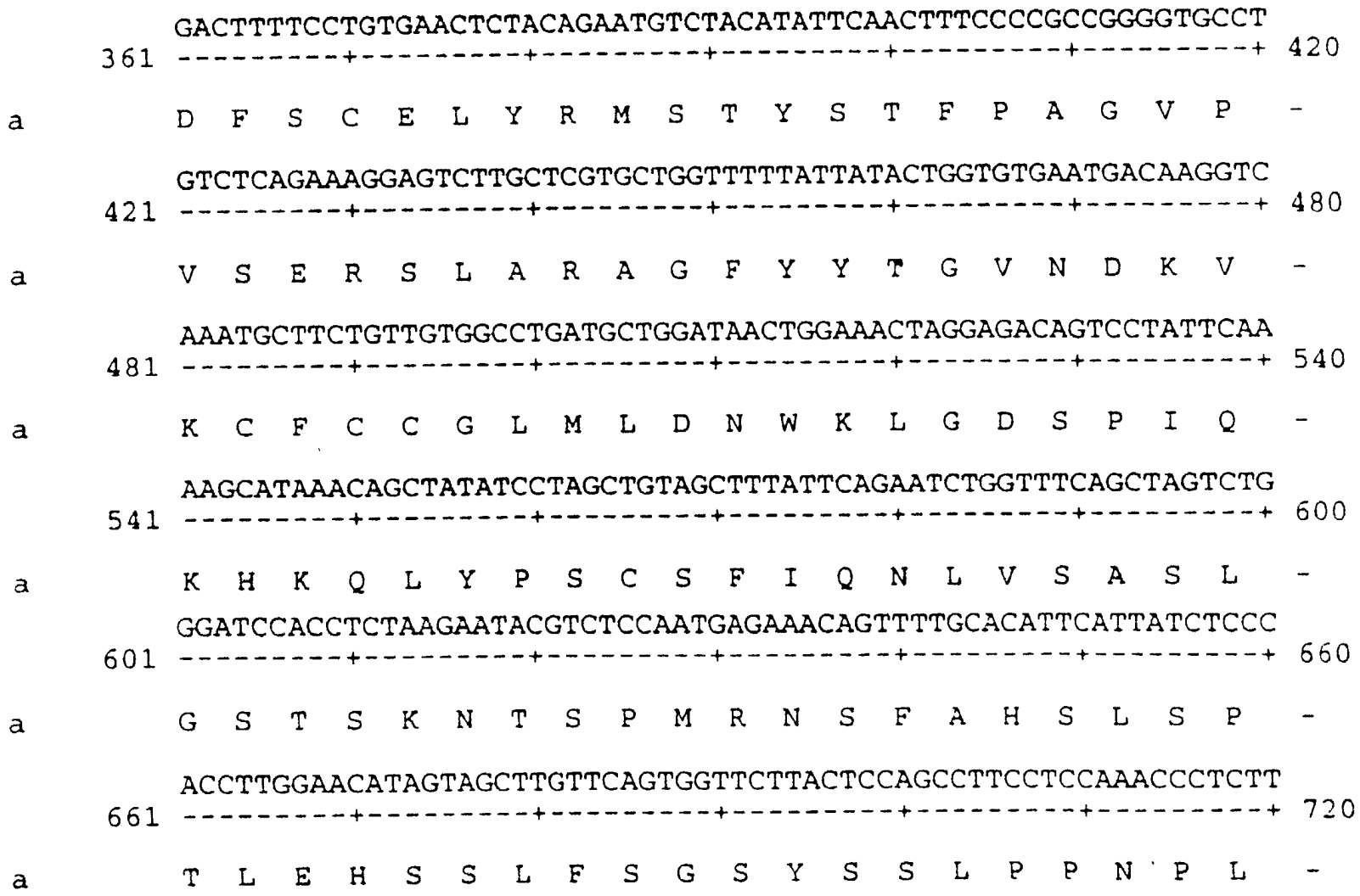

FIG. 3 (PAGE 2 OF 7)

HUMAN hiap-2

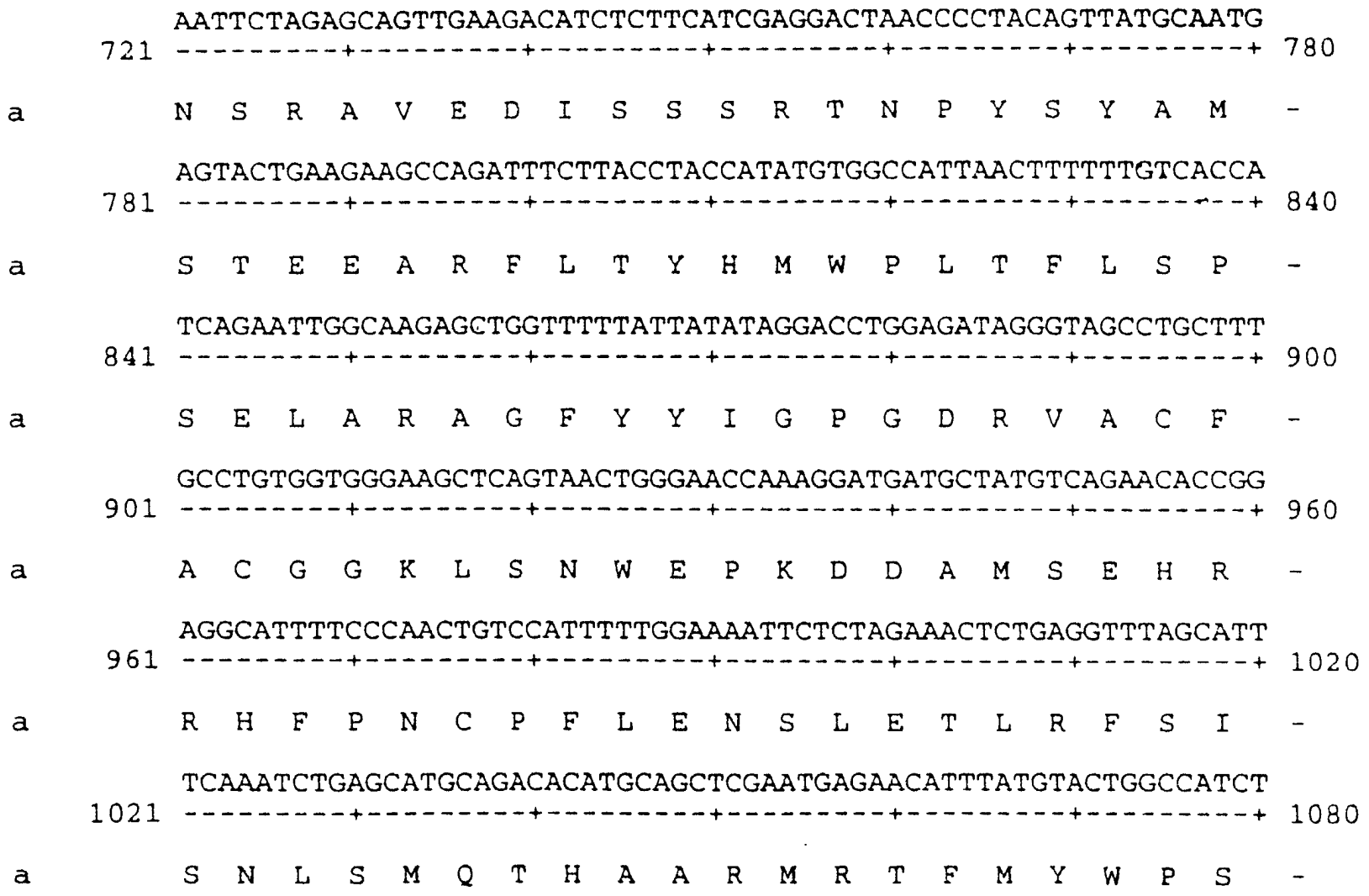

FIG. 3 (PAGE 3 OF 7)

HUMAN hiap-2

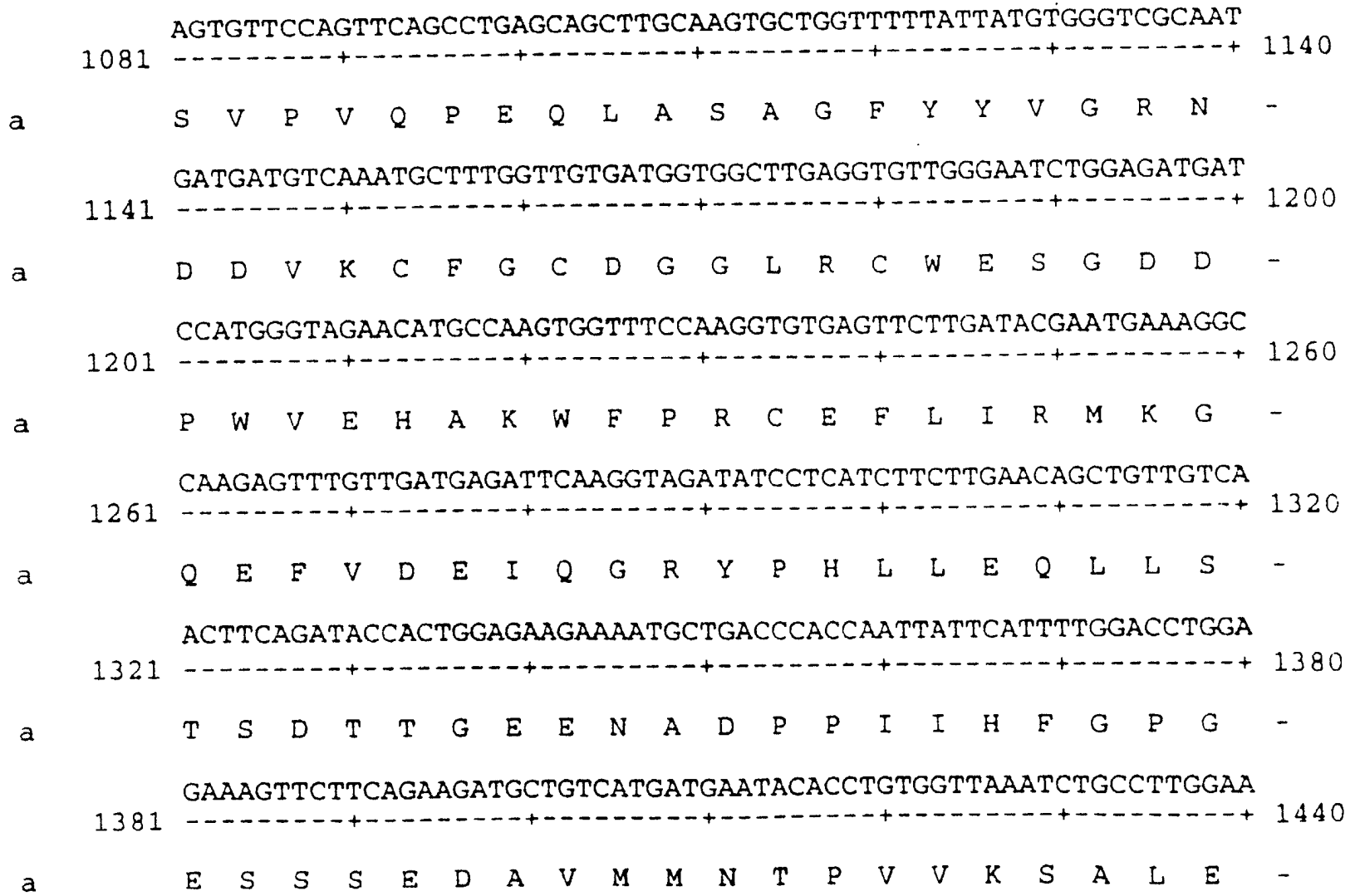

FIG. 3 (PAGE 4 OF 7)

HUMAN hiap-2

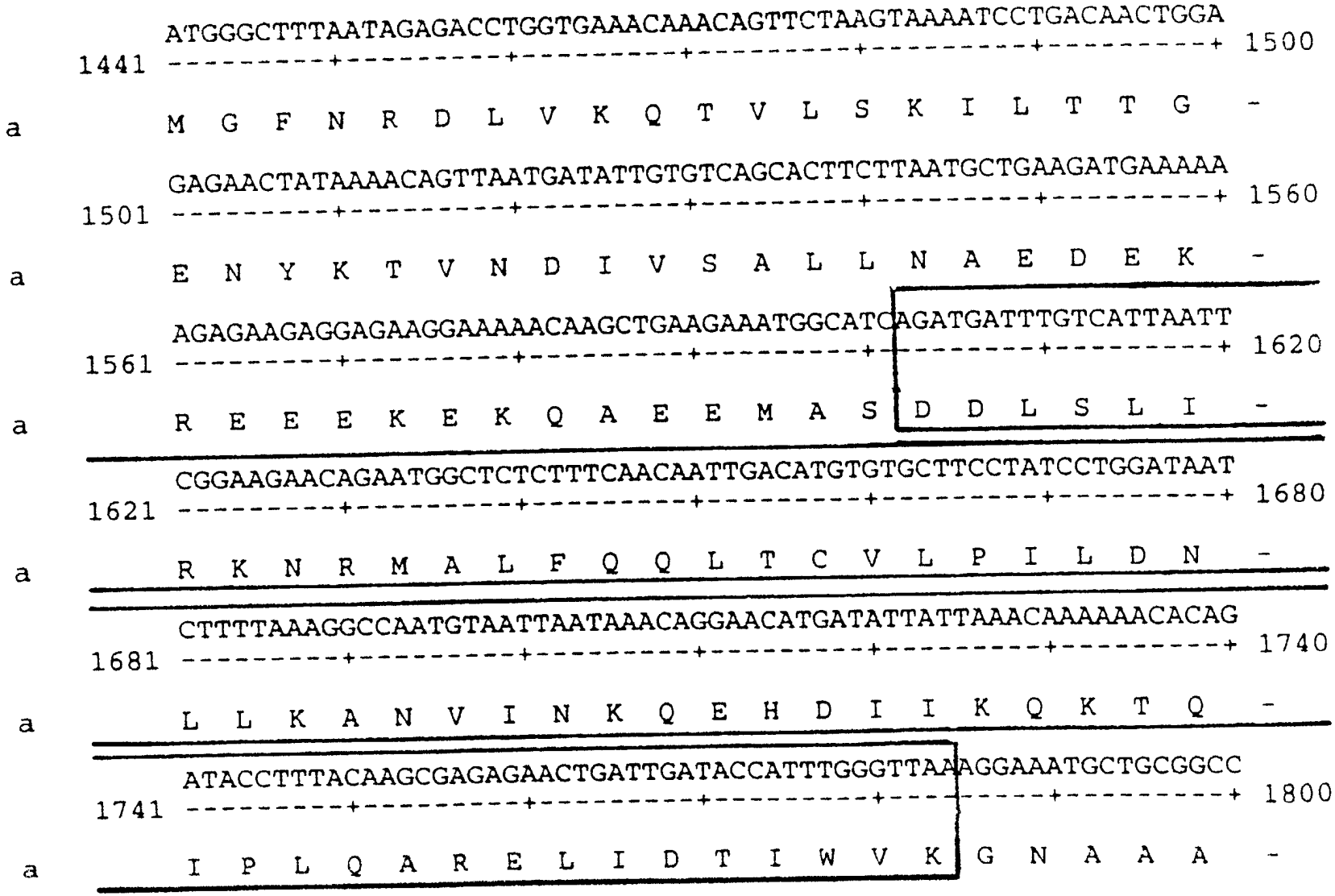

FIG. 3 (PAGE 5 OF 7)

HUMAN hiap-2

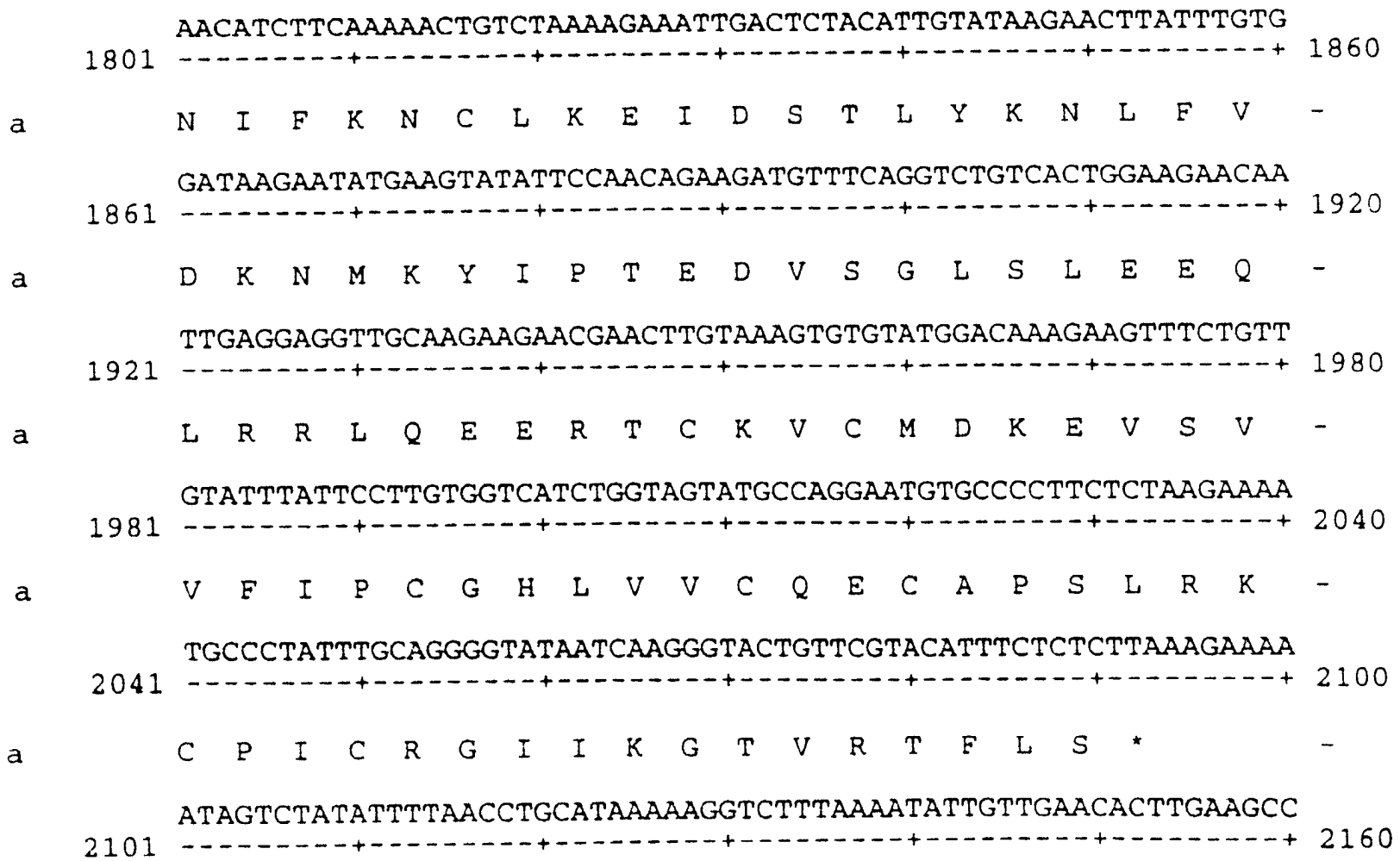

FIG. 3 (PAGE 6 OF 7)

HUMAN hiap-2

FIG. 3 (PAGE 7 OF 7)

MOUSE xiap

SEQ ID NO:9

```
          GACACTCTGCTGGGCGGCGGGCCGCCCTCCTCCGGGACCTCCCCTCGGGAACCGTCGCCC
        1 ---------+---------+---------+---------+---------+---------+ 60
a                                                                       -
          GCGGCGCTTAGTTAGGACTGGAGTGCTTGGCGCGAAAAGGTGGACAAGTCCTATTTTCCA
       61 ---------+---------+---------+---------+---------+---------+ 120
a                                                                       -
          GAGAAGATGACTTTTAACAGTTTTGAAGGAACTAGAACTTTTGTACTTGCAGACACCAAT
      121 ---------+---------+---------+---------+---------+---------+ 180
SEQ ID NO:10 a          M   T   F   N   S   F   E   G   T   R   T   F   V   L   A   D   T   N   -
          AAGGATGAAGAATTTGTAGAAGAGTTTAATAGATTAAAAACATTTGCTAACTTCCCAAGT
      181 ---------+---------+---------+---------+---------+---------+ 240
a         K   D   E   E   F   V   E   E   F   N   R   L   K   T   F   A   N   F   P   S   -
          AGTAGTCCTGTTTCAGCATCAACATTGGCGCGAGCTGGGTTTCTTTATACCGGTGAAGGA
      241 ---------+---------+---------+---------+---------+---------+ 300
a         S   S   P   V   S   A   S   T   L   A   R   A   G   F   L   Y   T   G   E   G   -
          GACACCGTGCAATGTTTCAGTTGTCATGCGGCAATAGATAGATGGCAGTATGGAGACTCA
      301 ---------+---------+---------+---------+---------+---------+ 360
a         D   T   V   Q   C   F   S   C   H   A   A   I   D   R   W   Q   Y   G   D   S   -
```

FIG. 4 (PAGE 1 OF 6)

MOUSE xiap

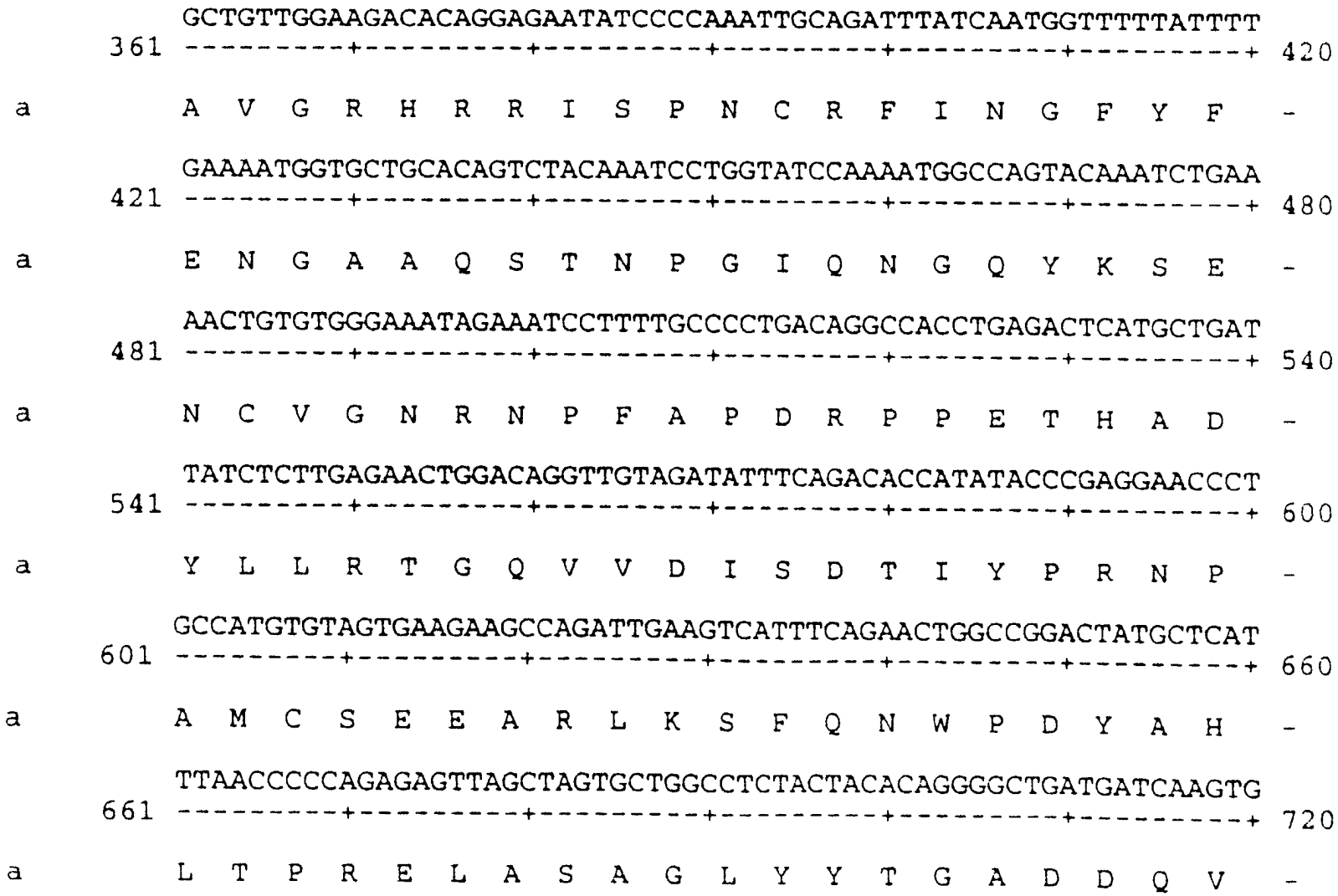

FIG. 4 (PAGE 2 OF 6)

MOUSE xiap

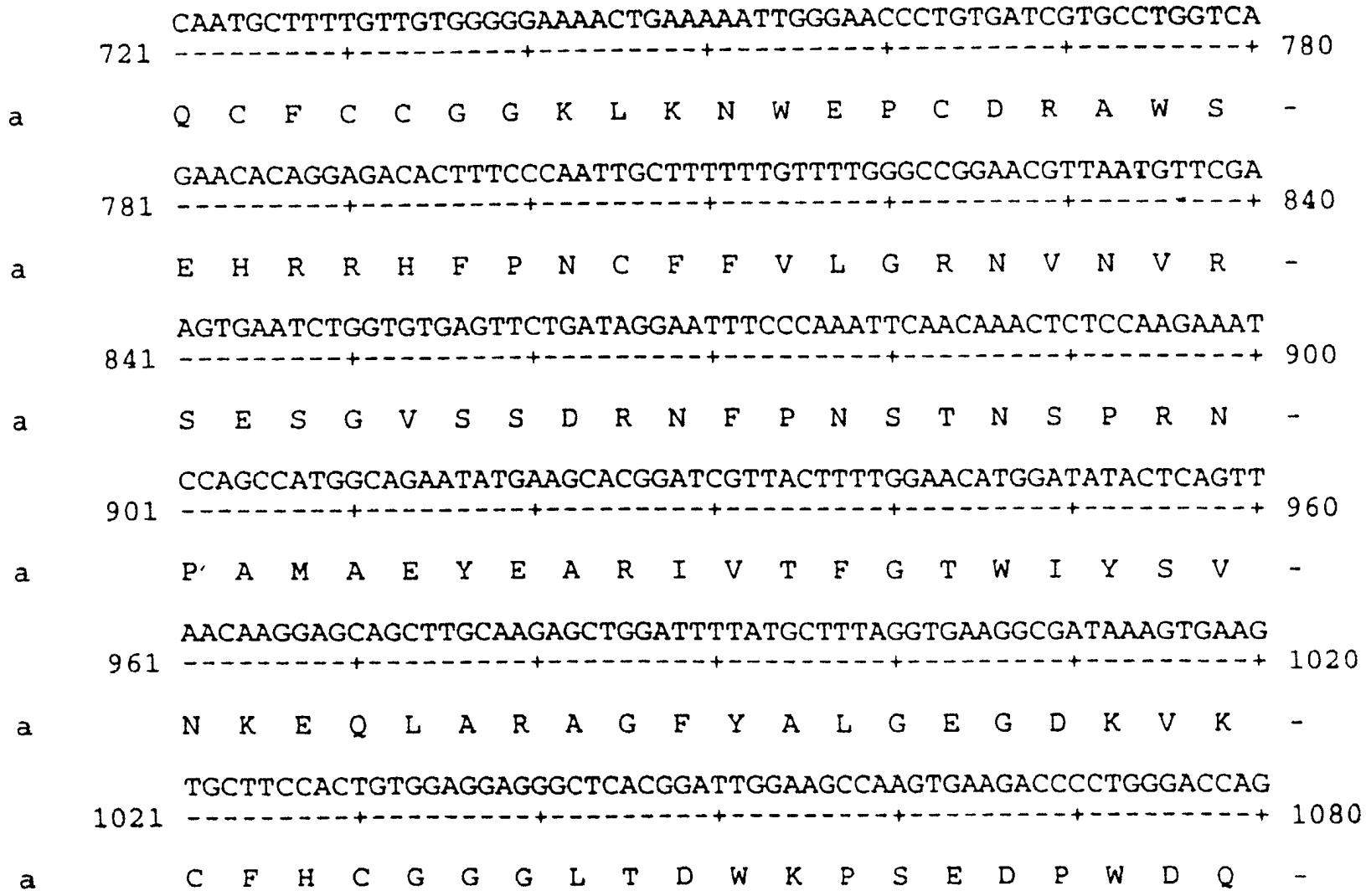

FIG. 4 (PAGE 3 OF 6)

MOUSE xiap

```
      CATGCTAAGTGCTACCCAGGGTGCAAATACCTATTGGATGAGAAGGGGCAAGAATATATA
1081  ---------+---------+---------+---------+---------+---------+ 1140

a       H  A  K  C  Y  P  G  C  K  Y  L  L  D  E  K  G  Q  E  Y  I  -

      AATAATATTCATTTAACCCATCCACTTGAGGAATCTTTGGGAAGAACTGCTGAAAAAACA
1141  ---------+---------+---------+---------+---------+---------+ 1200

a       N  N  I  H  L  T  H  P  L  E  E  S  L  G  R  T  A  E  K  T  -

      CCACCGCTAACTAAAAAAATCGATGATACCATCTTCCAGAATCCTATGGTGCAAGAAGCT
1201  ---------+---------+---------+---------+---------+---------+ 1260

a       P  P  L  T  K  K  I  D  D  T  I  F  Q  N  P  M  V  Q  E  A  -

      ATACGAATGGGATTTAGCTTCAAGGACCTTAAGAAAACAATGGAAGAAAAAATCCAAACA
1261  ---------+---------+---------+---------+---------+---------+ 1320

a       I  R  M  G  F  S  F  K  D  L  K  K  T  M  E  E  K  I  Q  T  -

      TCCGGGAGCAGCTATCTATCACTTGAGGTCCTGATTGCAGATCTTGTGAGTGCTCAGAAA
1321  ---------+---------+---------+---------+---------+---------+ 1380

a       S  G  S  S  Y  L  S  L  E  V  L  I  A  D  L  V  S  A  Q  K  -

      GATAATACGGAGGATGAGTCAAGTCAAACTTCATTGCAGAAAGACATTAGTACTGAAGAG
1381  ---------+---------+---------+---------+---------+---------+ 1440

a       D  N  T  E  D  E  S  S  Q  T  S  L  Q  K  D  I  S  T  E  E  -
```

FIG. 4 (PAGE 4 OF 6)

MOUSE xiap

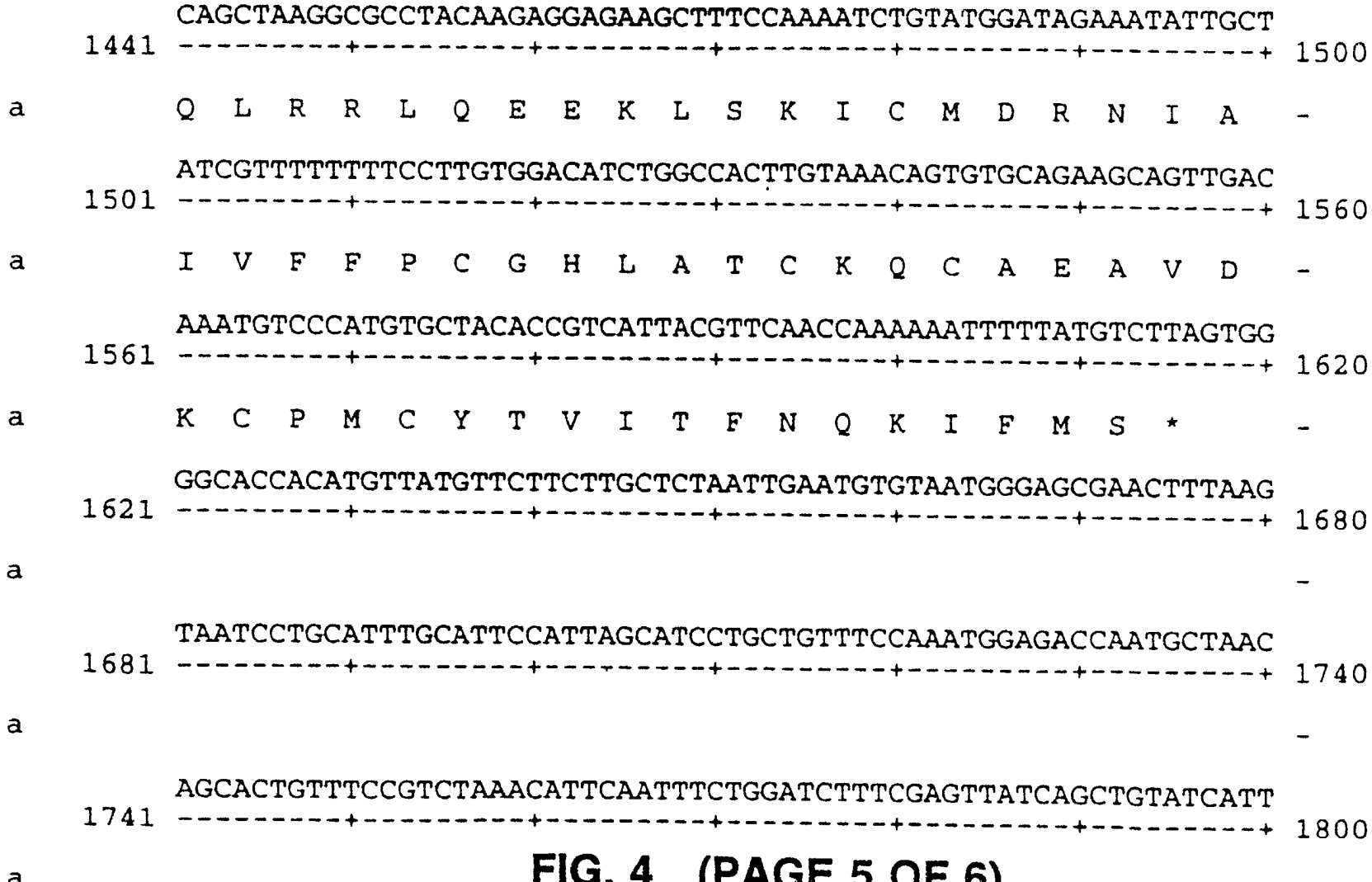

FIG. 4 (PAGE 5 OF 6)

MOUSE xiap

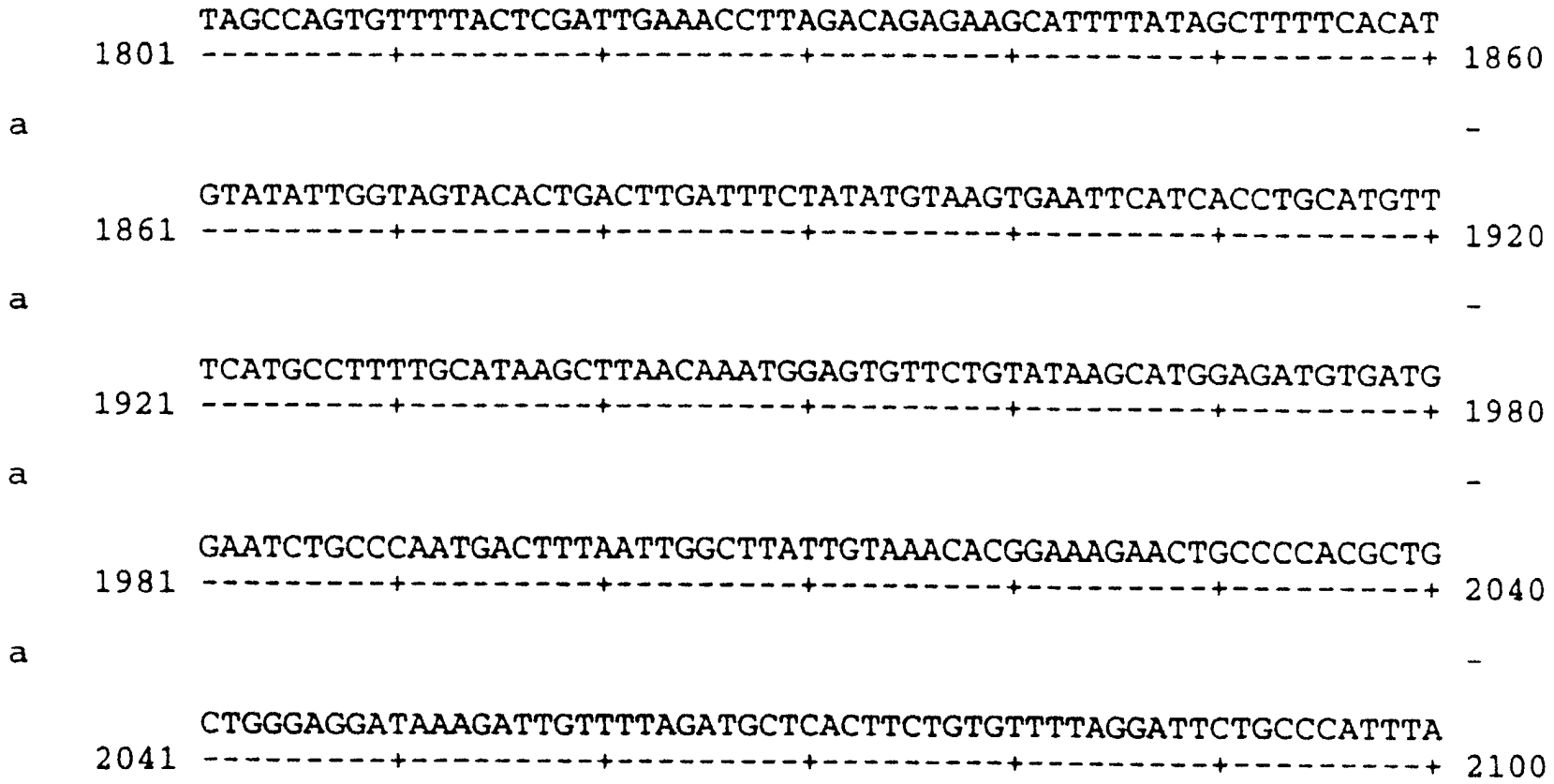

FIG. 4 (PAGE 6 OF 6)

M-hiap-1

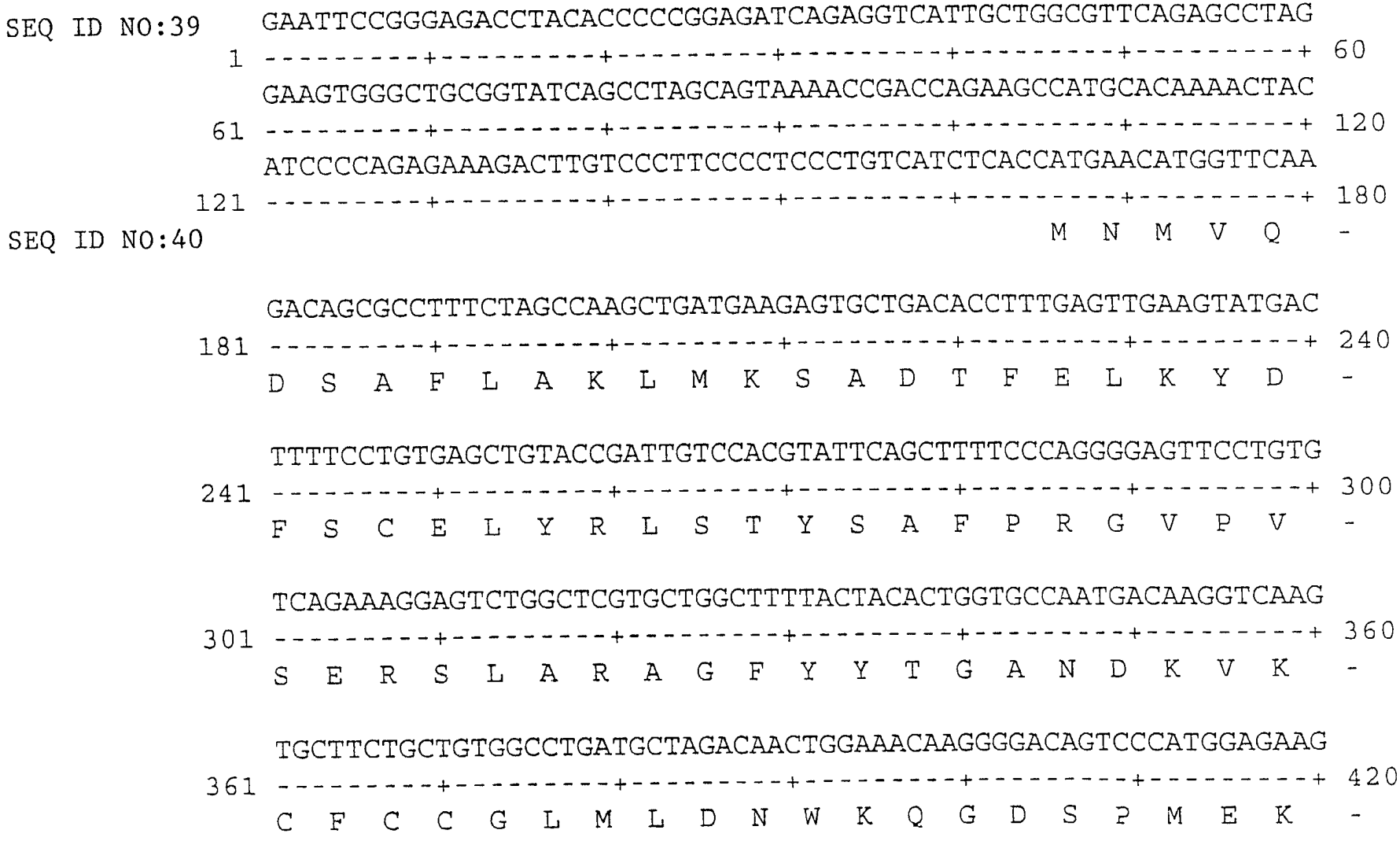

FIG. 5 (PAGE 1 OF 6)

M-hiap-1

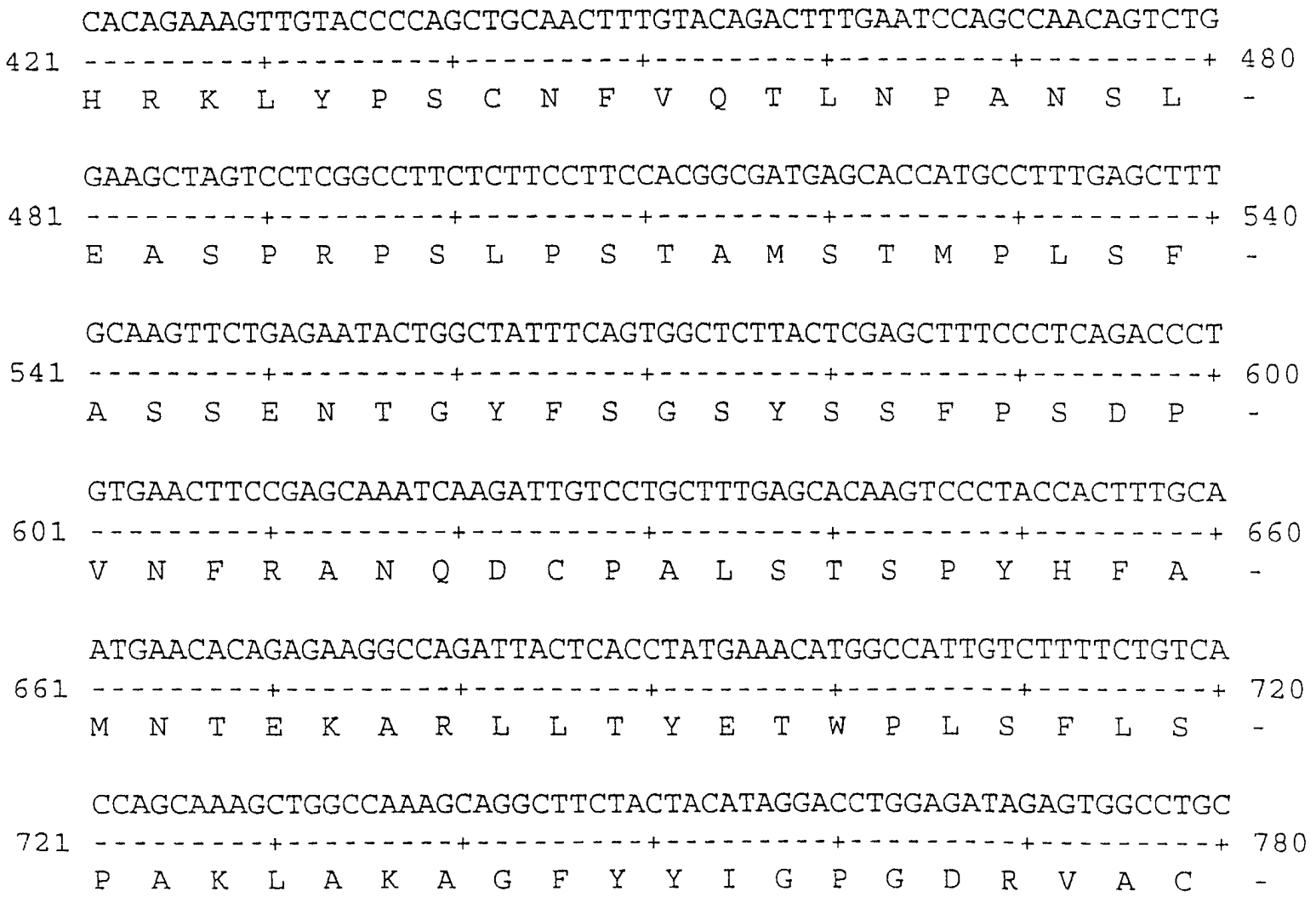

FIG. 5 (PAGE 2 OF 6)

M-hiap-1

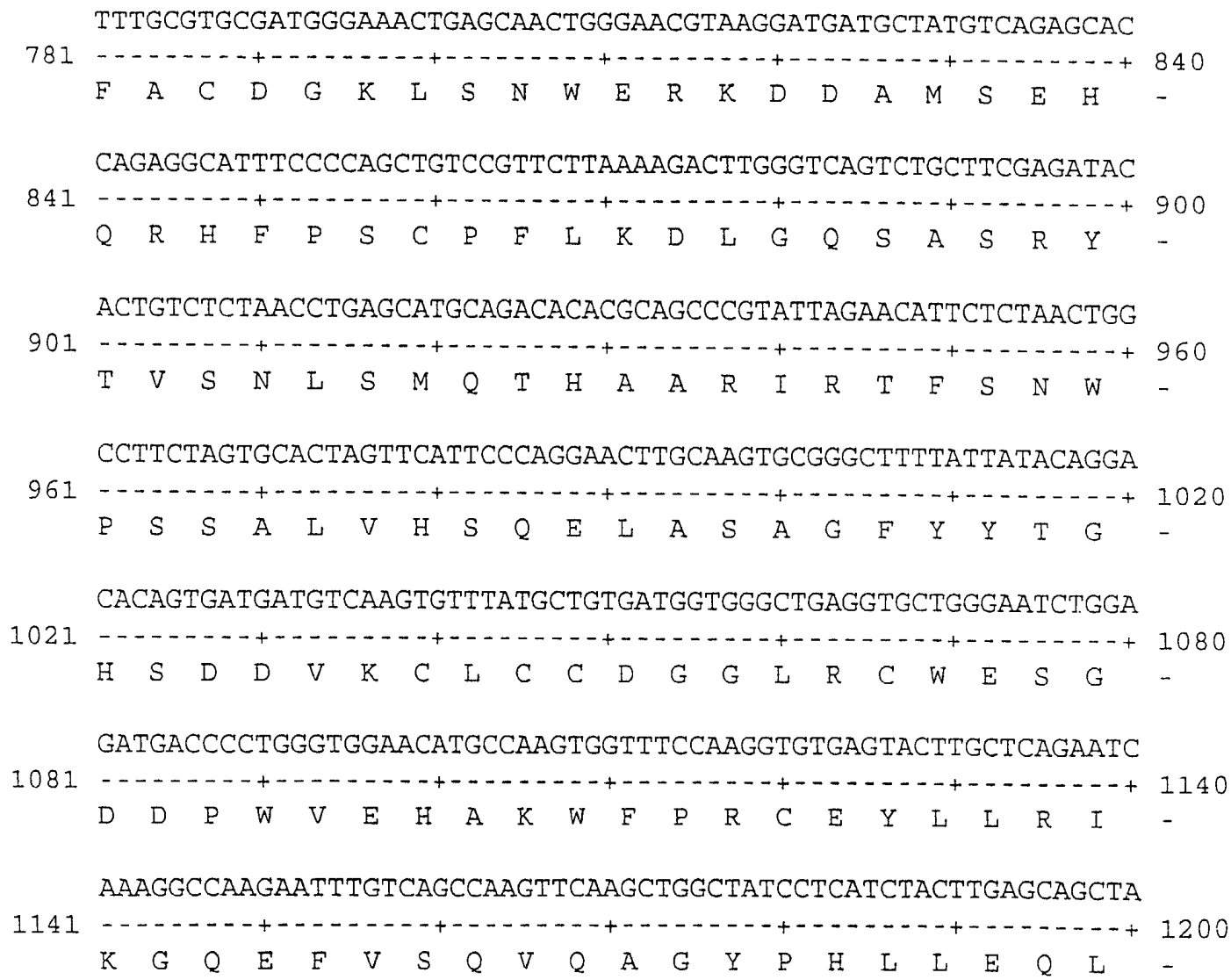

FIG. 5 (PAGE 3 OF 6)

M-hiap-1

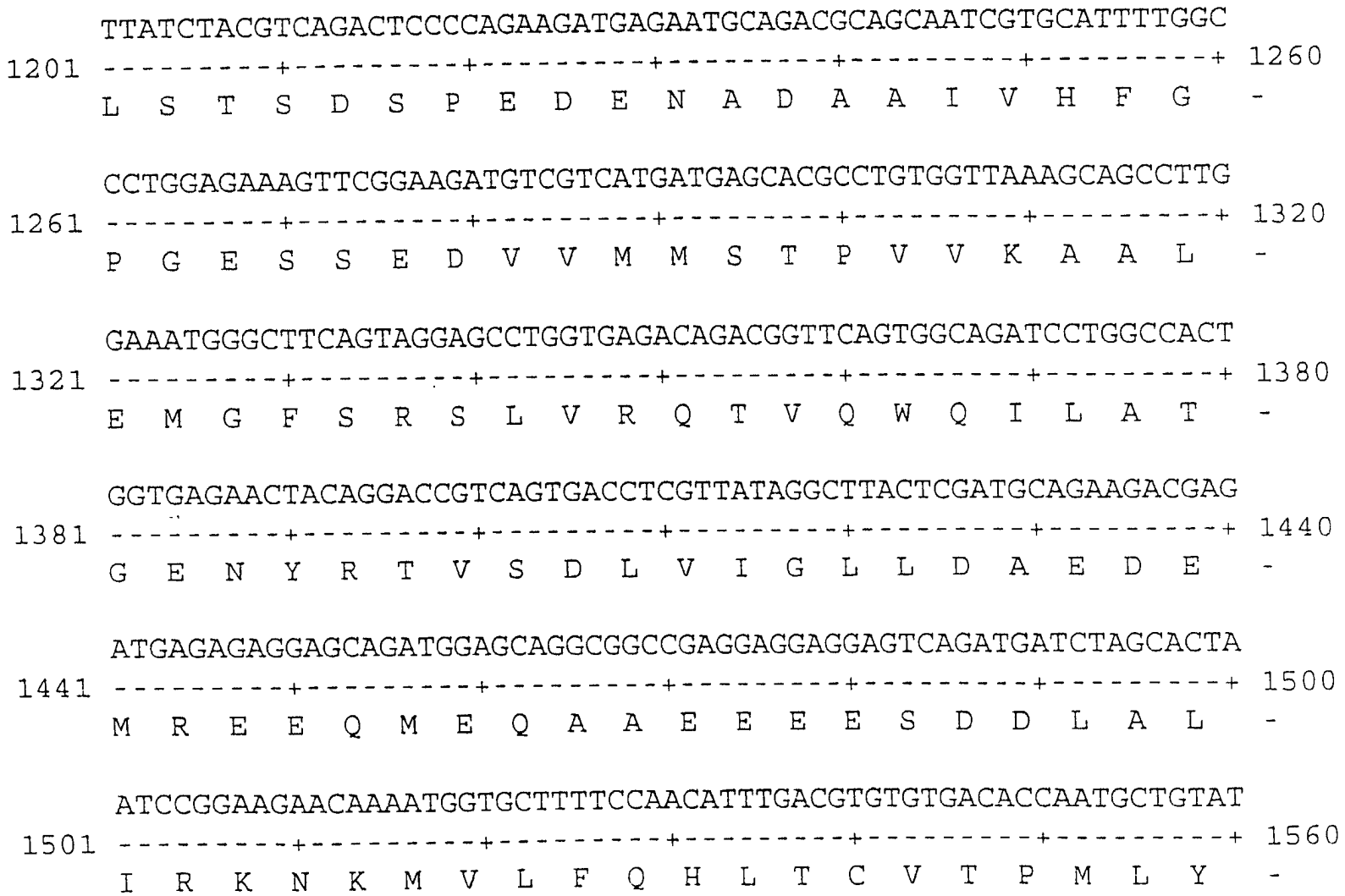

FIG. 5 (PAGE 4 OF 6)

M-hiap-1

```
     TGCCTCCTAAGTGCAAGGGCCATCACTGAACAGGAGTGCAATGCTGTGAAACAGAAACCA
1561 ---------+---------+---------+---------+---------+---------+ 1620
     C  L  L  S  A  R  A  I  T  E  Q  E  C  N  A  V  K  Q  K  P   -

     CACACCTTACAAGCAAGCACACTGATTGATACTGTGTTAGCAAAAGGAAACACTGCAGCA
1621 ---------+---------+---------+---------+---------+---------+ 1680
     H  T  L  Q  A  S  T  L  I  D  T  V  L  A  K  G  N  T  A  A   -

     ACCTCATTCAGAAACTCCCTTCGGGAAATTGACCCTGCGTTATACAGAGATATATTTGTG
1681 ---------+---------+---------+---------+---------+---------+ 1740
     T  S  F  R  N  S  L  R  E  I  D  P  A  L  Y  R  D  I  F  V   -

     CAACAGGACATTAGGAGTCTTCCCACAGATGACATTGCAGCTCTACCAATGGAAGAACAG
1741 ---------+---------+---------+---------+---------+---------+ 1800
     Q  Q  D  I  R  S  L  P  T  D  D  I  A  A  L  P  M  E  E  Q   -

     TTGCGGCCCCTCCCGGAGGACAGAATGTGTAAAGTGTGTATGGACCGAGAGGTATCCATC
1801 ---------+---------+---------+---------+---------+---------+ 1860
     L  R  P  L  P  E  D  R  M  C  K  V  C  M  D  R  E  V  S  I   -

     GTGTTCATTCCCTGTGGCCATCTGGTCGTGTGCAAAGACTGCGCTCCCTCTCTGAGGAAG
1861 ---------+---------+---------+---------+---------+---------+ 1920
     V  F  I  P  C  G  H  L  V  V  C  K  D  C  A  P  S  L  R  K   -
```

FIG. 5 (PAGE 5 OF 6)

M-hiap-1

```
     TGTCCCATCTGTAGAGGGACCATCAAGGGCACAGTGCGCACATTTCTCTCCTGAACAAGA
1921 ---------+---------+---------+---------+---------+---------+ 1980
     C  P  I  C  R  G  T  I  K  G  T  V  R  T  F  L  S  *        -

     CTAATGGTCCATGGCTGCAACTTCAGCCAGGAGGAAGTTCACTGTCACTCCCAGTTCCAT
1981 ---------+---------+---------+---------+---------+---------+ 2040
     TCGGAACTTGAGGCCAGCCTGGATAGCACGAGACACCGCCAAACACACAAATATAAACAT
2041 ---------+---------+---------+---------+---------+---------+ 2100
     GAAAAACTTTTGTCTGAAGTCAAGAATGAATGAATTACTTATATAATAATTTTAATTGGT
2101 ---------+---------+---------+---------+---------+---------+ 2160
     TTCCTTAAAAGTGCTATTTGTTCCCAACTCAGAAAATTGTTTTCTGTAAACATATTTACA
2161 ---------+---------+---------+---------+---------+---------+ 2220
     TACTACCTGCATCTAAAGTATTCATATATTCATATATTCAGATGTCATGAGAGAGGGTTT
2221 ---------+---------+---------+---------+---------+---------+ 2280
     TGTTCTTGTTCCTGAAAAGCTGGTTTATCATCTGATCAGCATATACTGCGCAACGGGCAG
2281 ---------+---------+---------+---------+---------+---------+ 2340
     GGCTAGAATCCATGAACCAAGCTGCAAAGATCTCACGCTAAATAAGGCGGAAAGATTTGG
2341 ---------+---------+---------+---------+---------+---------+ 2400
     AGAAACGAAAGGAAATTCTTTCCTGTCCAATGTATACTCTTCAGACTAATGACCTCTTCC
2401 ---------+---------+---------+---------+---------+---------+ 2460
     TATCAAGCCTTCTA
2461 ----------+---- 2474
```

FIG. 5 (PAGE 6 OF 6)

M-hiap-2

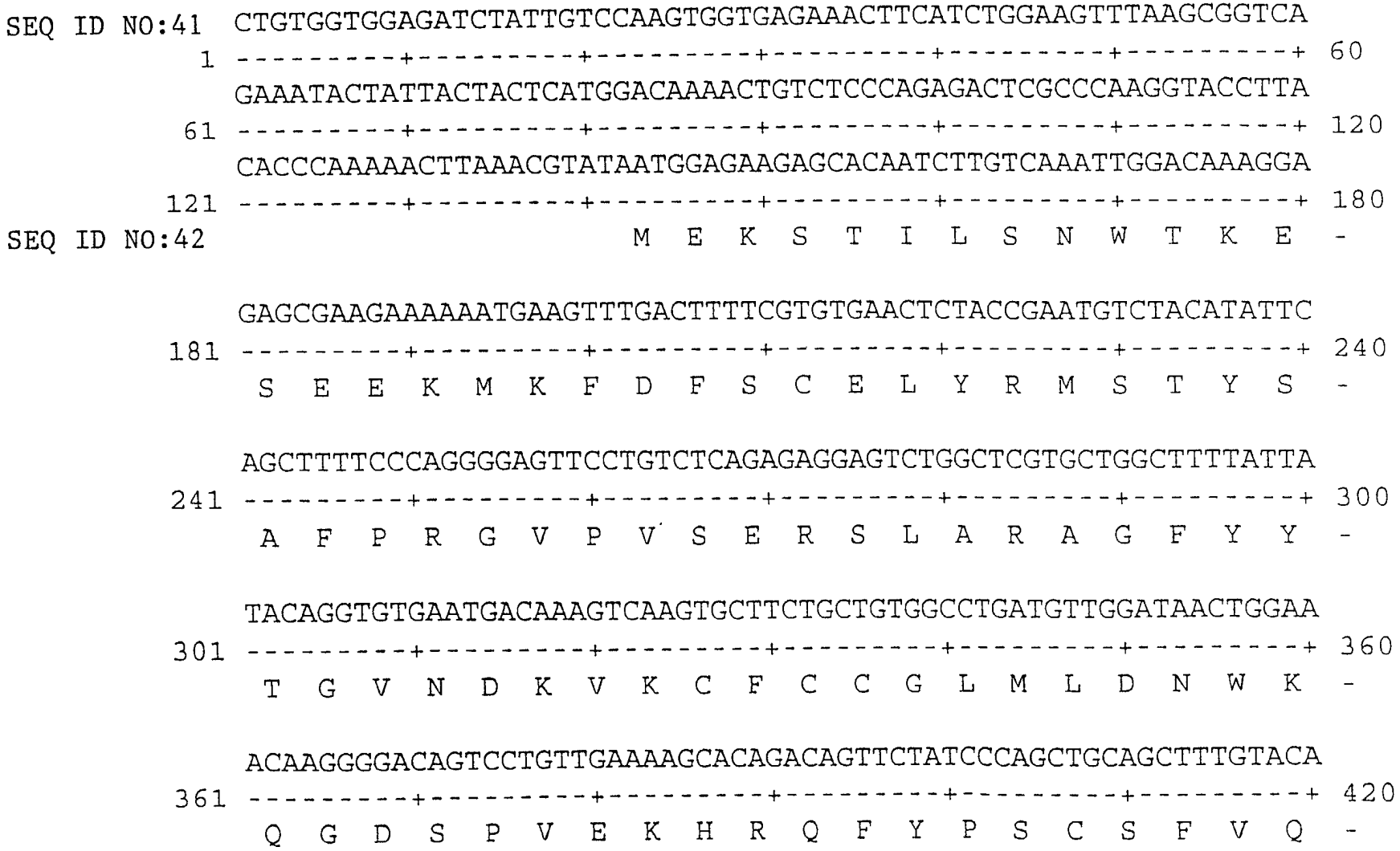

FIG. 6 (PAGE 1 OF 6)

M-hiap-2

```
    GACTCTGCTTTCAGCCAGTCTGCAGTCTCCATCTAAGAATATGTCTCCTGTGAAAAGTAG
421 ---------+---------+---------+---------+---------+---------+ 480
     T  L  L  S  A  S  L  Q  S  P  S  K  N  M  S  P  V  K  S  R  -

    ATTTGCACATTCGTCACCTCTGGAACGAGGTGGCATTCACTCCAACCTGTGCTCTAGCCC
481 ---------+---------+---------+---------+---------+---------+ 540
     F  A  H  S  S  P  L  E  R  G  G  I  H  S  N  L  C  S  S  P  -

    TCTTAATTCTAGAGCAGTGGAAGACTTCTCATCAAGGATGGATCCCTGCAGCTATGCCAT
541 ---------+---------+---------+---------+---------+---------+ 600
     L  N  S  R  A  V  E  D  F  S  S  R  M  D  P  C  S  Y  A  M  -

    GAGTACAGAAGAGGCCAGATTTCTTACTTACAGTATGTGGCCTTTAAGTTTTCTGTCACC
601 ---------+---------+---------+---------+---------+---------+ 660
     S  T  E  E  A  R  F  L  T  Y  S  M  W  P  L  S  F  L  S  P  -

    AGCAGAGCTGGCCAGAGCTGGCTTCTATTACATAGGGCCTGGAGACAGGGTGGCCTGTTT
661 ---------+---------+---------+---------+---------+---------+ 720
     A  E  L  A  R  A  G  F  Y  Y  I  G  P  G  D  R  V  A  C  F  -

    TGCCTGTGGTGGGAAACTGAGCAACTGGGAACCAAAGGATTATGCTATGTCAGAGCACCG
721 ---------+---------+---------+---------+---------+---------+ 780
     A  C  G  G  K  L  S  N  W  E  P  K  D  Y  A  M  S  E  H  R  -
```

FIG. 6 (PAGE 2 OF 6)

M-hiap-2

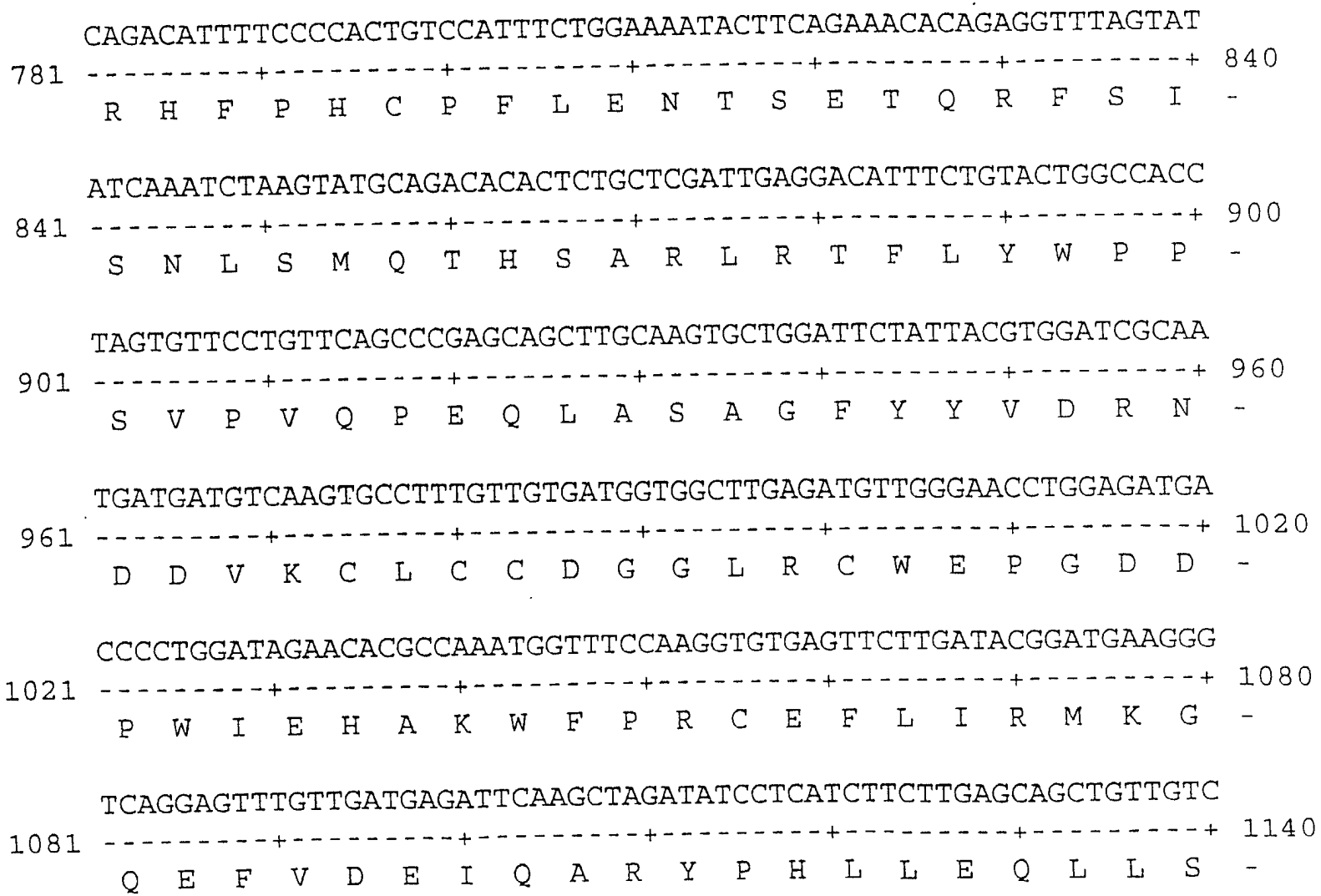

FIG. 6 (PAGE 3 OF 6)

M-hiap-2

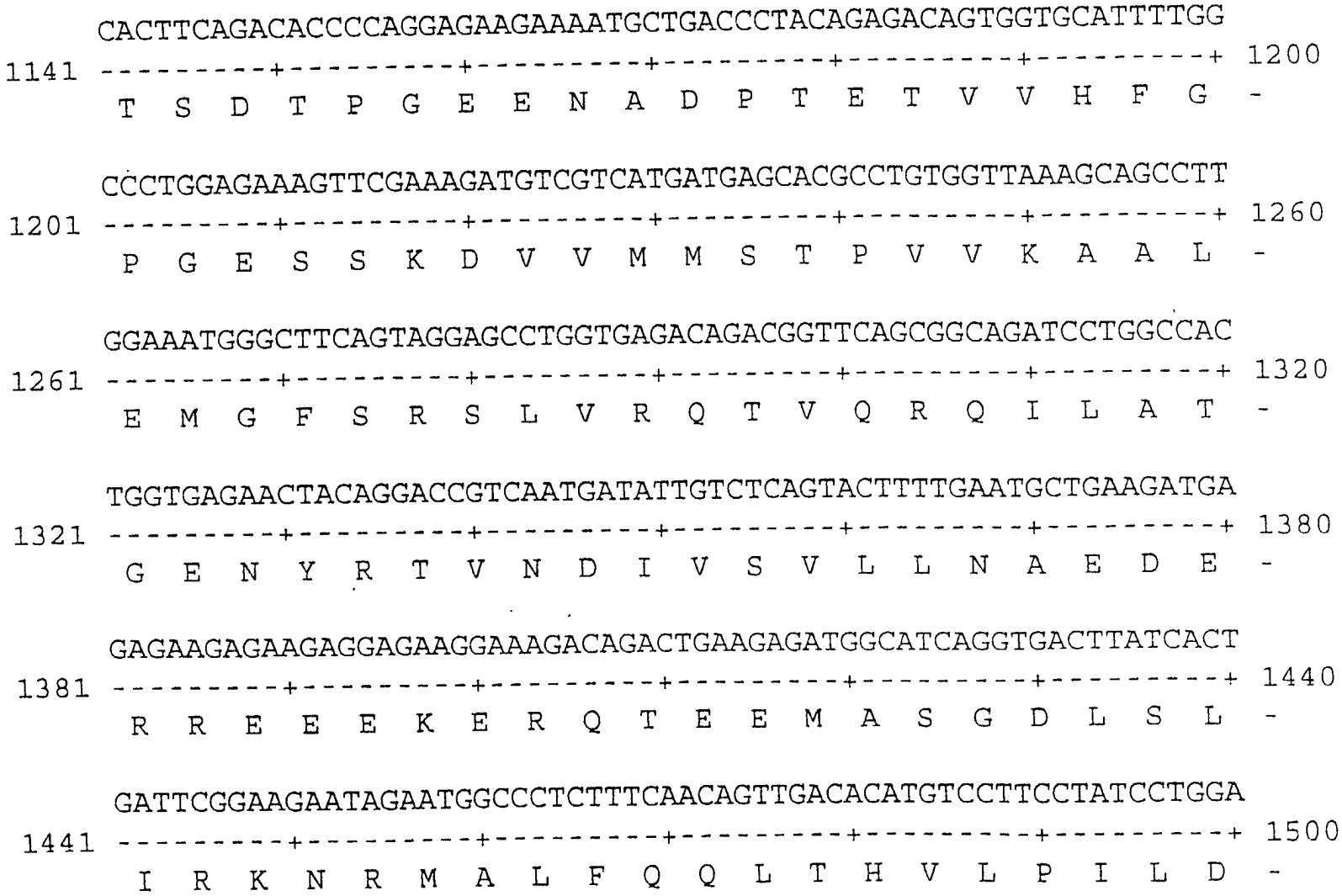

FIG. 6 (PAGE 4 OF 6)

M-hiap-2

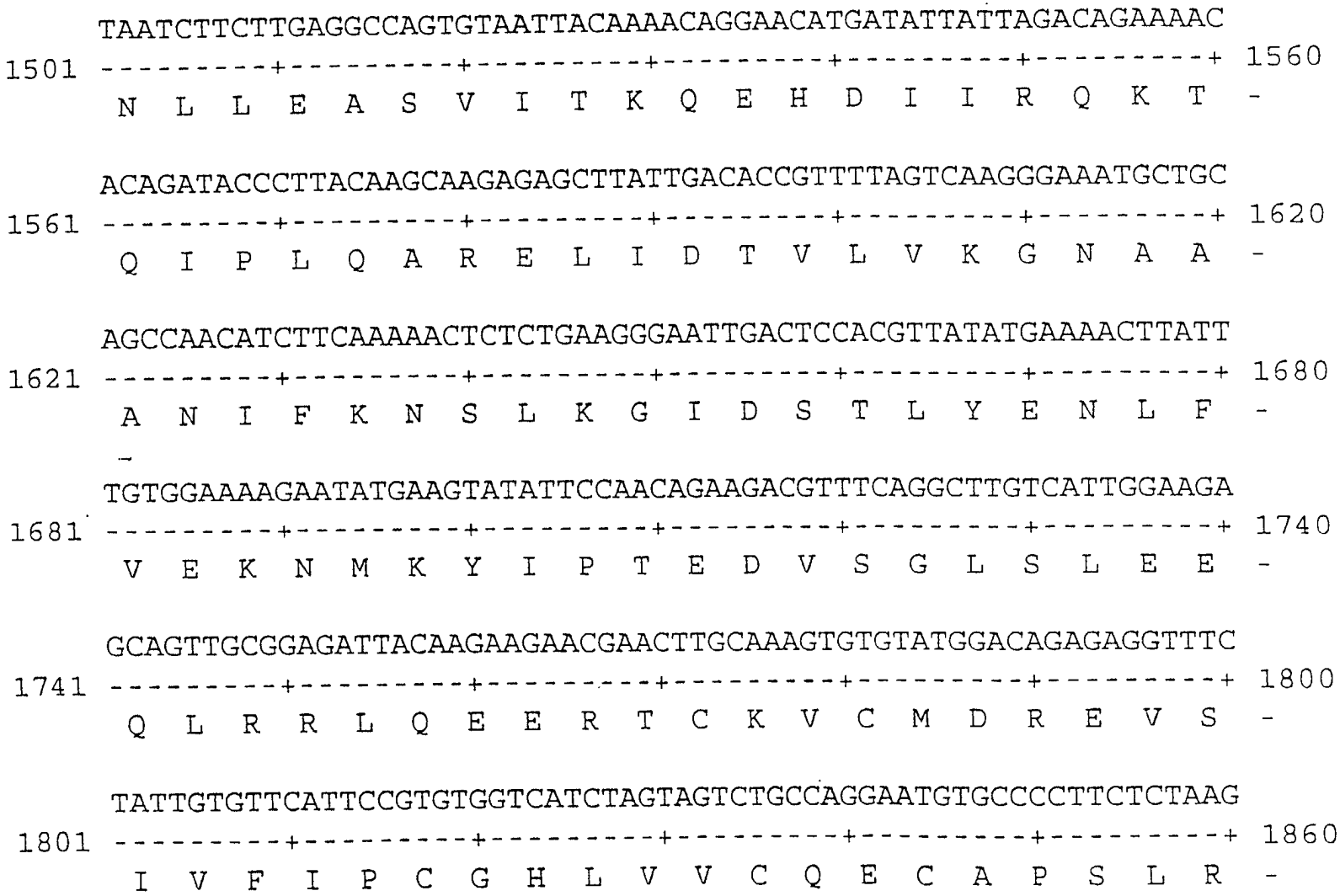

FIG. 6 (PAGE 5 OF 6)

M-hiap-2

```
     GAAGTGCCCCATCTGCAGGGGGACAATCAAGGGGACTGTGCGCACATTTCTCTCATGAGT
1861 ---------+---------+---------+---------+---------+---------+ 1920
      K  C  P  I  C  R  G  T  I  K  G  T  V  R  T  F  L  S  *     -

     GAAGAATGGTCTGAAAGTATTGTTGGACATCAGAAGCTGTCAGAACAAAGAATGAACTAC
1921 ---------+---------+---------+---------+---------+---------+ 1980
     TGATTTCAGCTCTTCAGCAGGACATTCTACTCTCTTTCAAGATTAGTAATCTTGCTTTAT
1981 ---------+---------+---------+---------+---------+---------+ 2040
     GAAGGGTAGCATTGTATATTTAAGCTTAGTCTGTTGCAAGGGAAGGTCTATGCTGTTGAG
2041 ---------+---------+---------+---------+---------+---------+ 2100
     CTACAGGACTGTGTCTGTTCCAGAGCAGGAGTTGGGATGCTTGCTGTATGTCCTTCAGGA
2101 ---------+---------+---------+---------+---------+---------+ 2160
     CTTCTTGGGATTTGGGAATTTGGGGAAAGCTTTGGAATCCAGTGATGTGGAGCTCAGAAA
2161 ---------+---------+---------+---------+---------+---------+ 2220
     TCCTGGAACCAGTGACTCTGGTACTCAGTAGATAGGGTACCCTGTACTTCTTGGTGCTTT
2221 ---------+---------+---------+---------+---------+---------+ 2280
     TCCAGTCTGGGAAATAAGGAGGAATCTGCTGCTGGTAAAAATTTGCTGGATGTGAGAAAT
2281 ---------+---------+---------+---------+---------+---------+ 2340
     AGATGAAAGTGTTTCGGGTGGGGGCGTGCATCAGTGTAGTGTGTGCAGGGATGTATGCAG
2341 ---------+---------+---------+---------+---------+---------+ 2400
     GCCAAACACTGTGTAG
2401 ---------+------ 2416
```

FIG. 6 (PAGE 6 OF 6)

Alignment of BIR (Baculoviral IAP Repeats) Domains

Baculovirus

Cp_iap	Cydia pomonella
Op_iap	Orgyia pseudotsugata

Human

xiap	IAP on X chromosome
hiap1, hiap2	two different human IAP genes

Mouse

m-xiap	mouse homologue of human xiap gene

Insect

diap	Drosophilia IAP gene, not clearly a homologue of xiap or hiap

note on consensus: The consensus line represents amino acids or very similar amino acids which are present in 14 of the 19 BIR sequences at each position. Capitalized residues are those that are in the consensus sequence.

FIG. 7

```
                                  1                                                  68
SEQ ID NO:11   Op_iap-1     kaaRLgTYtn WPvqf.leps rMAasGFYY1 GrgDeVrCaf CkveitnWvr gDdpeLdHkr waPqCpFV
SEQ ID NO:14   Cp_iap-1     eevRLnTFek WPvsf.lspe tMAknGFYY1 GrsDeVrCaf CkveimrWke gEdpaadHkk waPqCpFV
SEQ ID NO:15   diap-2       eanRLvTFkd WPnpn.iLpq aLAkAGFYY1 nrlDhVkCvw CngviakWek nDnafeeHkr ffPqCprV
SEQ ID NO:16   m-xiap-1     efnRLkTFan FPsspvsas tLArAGFLYt GegDtVqCFs ChaaidrWqy gDsavgrHrr isPnCrFI
SEQ ID NO:17   xiap-1       efnRLkTFan FPsgspvsas tLArAGFLYt GegDtVrCFs ChaavdrWqy gDsavgrHrk vsPnCrFI
SEQ ID NO:18   hiap1-1      elyRMsTYst FPagvpvser sLArAGFYYt GvnDkVkCFc CglmldnWkr gDsptekHkk lyPsCrFV
SEQ ID NO:19   hiap2-1      elyRMsTYst FPagvpvser sLArAGFYYt GvnDkVkCFc CglmldnWkl gDspiqkHkq lyPsCsFI
SEQ ID NO:20   m-xiap-2     eeaRLksFqn WPdyahltpr eLAsAGLYYt GadDqVqCFc CggklknWep cDrawseHrr hfPnCfFV
SEQ ID NO:21   xiap-2       eeaRLksFqn WPdyahltpr eLAsAGLYYt GigDqVqCFc CggklknWep cDrawseHrr hfPnCfFV
SEQ ID NO:22   hiap1-2      enaRLlTFqt WP.lLflspt dLArAGFYYi GpgDrVaCFa CggklsnWep kDnamseHlr hfPkCpFI
SEQ ID NO:23   hiap2-2      eeaRFlTYhm WP.lLflsps eLArAGFYYi GpgDrVaCFa CggklsnWep kDdamseHrr hfPnCpF1
SEQ ID NO:24   m-xiap-3     yeaRivTFgt Wiysv..nke qLArAGFYal GegDkVkCFh CgggltdWkp sEdpwdqHak cyPgCkY1
SEQ ID NO:25   xiap-3       yeaRifTFgt Wiysv..nke qLArAGFYal GegDkVkCFh CgggltdWkp sEdpweqHak wyPgCkY1
SEQ ID NO:26   hiap1-3      haaRFkTFfn WPssvlvnpe qLAsAGFYYv GnsDdVkCFc CdgglrcWes gDdpwvqHak wfPrCeY1
SEQ ID NO:27   hiap2-3      haaRMrTFmy WPssvpvqpe qLAsAGFYYv GrnDdVkCFg CdgglrcWes gDdpwveHak wfPrCeF1
SEQ ID NO:28   Op_iap-2     eaaRLrTFae WPrglkqrpe eLAeAGFFYt GqgDktrCFc CdgglkdWep dDapwqqHar wydrCeYV
SEQ ID NO:29   Cp_iap-2     eaaRvksFhn WPrcmkqrpe qMAdAGFFYt GygDntkCFy CdgglkdWep eDvpweqHvr widrCaYV
SEQ ID NO:30   diap-3       vdaRLrTFtd WPisniqpas aLAqAGLYYq kigDqVrCFh CniglrsWqk eDepwieHak wsPkCqFV
SEQ ID NO:31   diap-1       esvRLaTFge WPlnapvsae dLvanGFF.. GtwmeaeCdf ChvridrWey gDlvaerHrr ssPiCsmV
SEQ ID NO:2    Consensus    ---RL-TF-- WP-------- -LA-AGFYY- G--D-V-CF- C------W-- -D-----H-- --P-C-FV
```

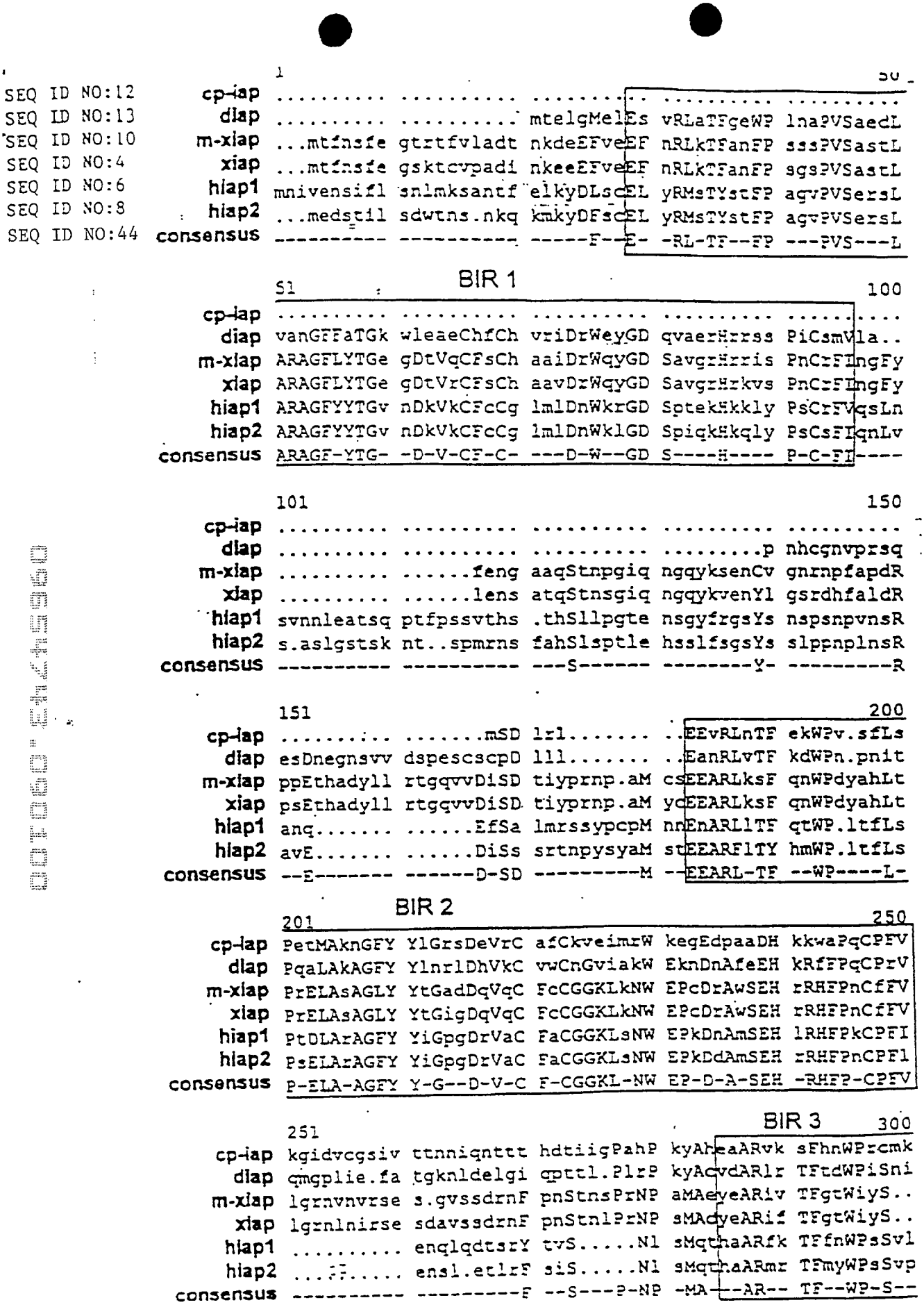

Fig. 8 (page 1 of 3)

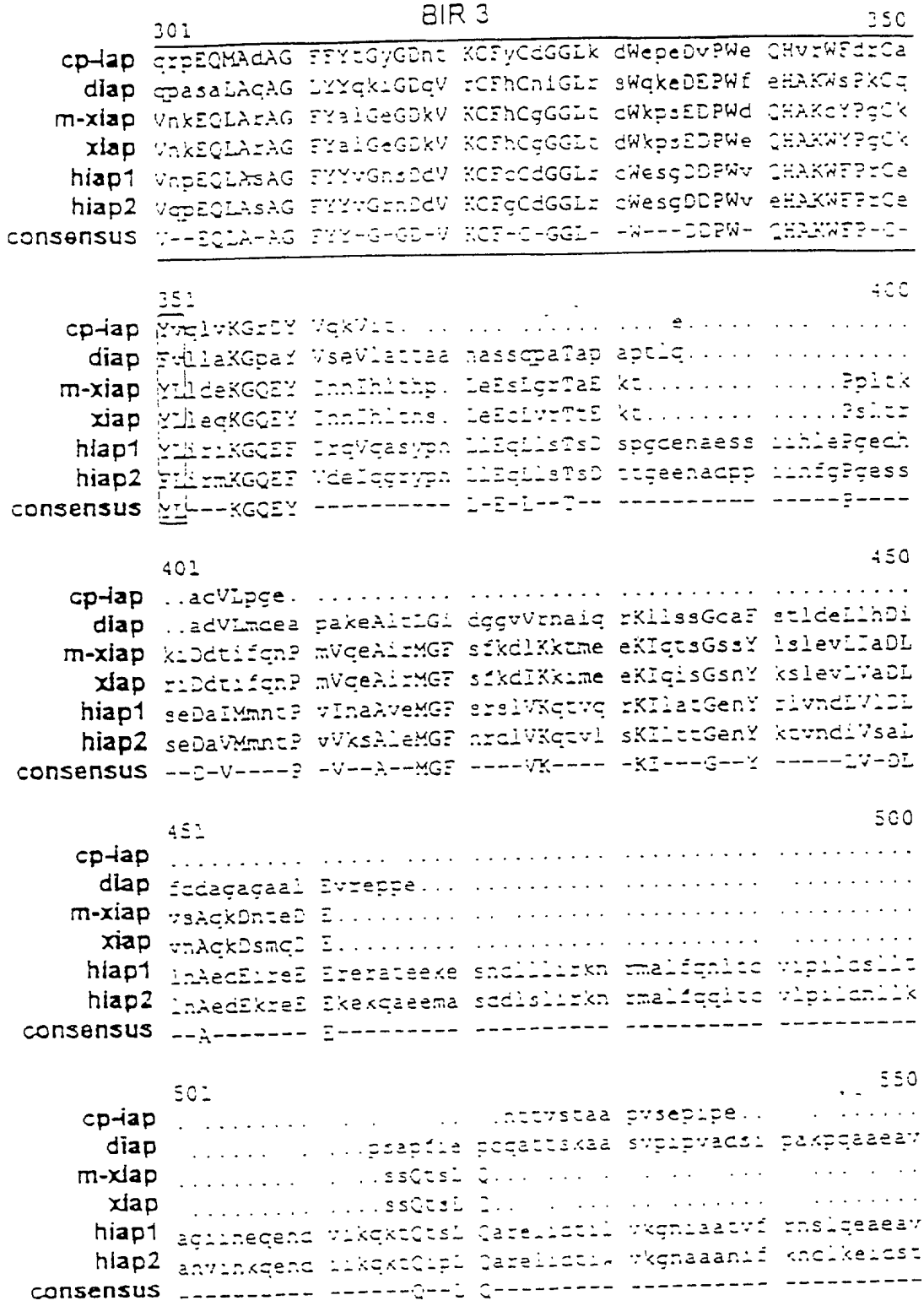

Fig. 8 (page 2 of 3)

```
                                              Ring Zinc Finger
          551                                                 600
cp-iap    ...tki.... .......... .....Ekepq veDskLCKIC yveEciVcFV
diap      sniskitdei qkmsvstpng nlSlEEenRq LkDarLCKVC LDeEVqVVFl
m-xiap    .......... .........k aiStEEQLRR LqEEkLsKIC MDrnIaIVFf
xiap      .......... .........k eiStEEQLRR LqEEkLCKIC MDrnIaIVFV
hiap1     lvehlfvqqd ikyiptedvs dlpvEEQLRR LpEErtCKVC MDkEVsIVFI
hiap2     lyknlfvdkn mkyiptedvs glSlEEQLRR LqEErtCKVC MDkEVsVVFI
consensus ---------- ---------- --S-EEQLRR L-EE-LCK-C MD-EV--VF-

          601                               635
cp-iap    PCGHvVaCak CAlSVdKCPM CRkIVtsvlk vYFS.
diap      PCGHLatCnq CApSVanCPM CRadIkqfvr tFLS*
m-xiap    PCGHLatCkq CAeaVdKCPM CytVItfnqk iFMS*
xiap      PCGHLVtCkq CAeaVdKCPM CytVItfkqk iFMS*
hiap1     PCGHLVvCkd CApSlrKCPi CRstIkqtvr tFLS*
hiap2     PCGHLVvCqe CApSlrKCPi CRgIIkqtvr tFLS.
consensus PCGHLV-C-- CA-SV-KCPM CR--I----- -FLS-
```

Fig. 8 (page 3 of 3)

Alignment of RZF (Ring Zinc Finger) Domains

Baculovirus

Cp_iap	Cydia pomonella
Op_ap	Orgyia pseudotsugata

Human

xiap	IAP on X chromosome
hiap1, hiap2	two different human IAP genes

Mouse

m-xiap	mouse homologue of human xiap gene

Insect

diap	Drosophilia IAP gene, not clearly a homologue of xiap or hiap

note on consensus: The consensus line represents amino acids or very similar amino acids which are present in 6 of the 7 RZF sequences at each position. Capitalized residues are those that are in the consensus sequence.

FIG. 9

```
                                1                                      46
SEQ ID NO:32  hiap2             EqlrrlqEer tCKVCMdkev sVvFiPCGH1 vvCqeCApel rkCPiC
SEQ ID NO:33  hiap1             EqltrlpEer tCKVCMdkev sIvFiPCGH1 w CkdCApsl rkCPiC
SEQ ID NO:34  m-xiap            EqltrlqEek lsKICMdrni aIvFfPCGH1 atCkqCAeav dkCPmC
SEQ ID NO:35  xiap              EqltrlqEek lCKICMdrni aIvFvPCGH1 vtCkqCAeav dkCPmC
SEQ ID NO:36  diap              EenrglkDar lCKVCLdeev gVvFlPCGH1 atCnqCApev anCPmC
SEQ ID NO:37  Cp_iap            EkepqveDsk lCKICyveec iVcFvPCGHv vaCakCAlsv dkCPmC
SEQ ID NO:38  Op_iap            aveaevaDdr lCKICLgack tVcFvPCGHv vaCgkCAagv ttCPvC
SEQ ID NO: 1  Consensus         E------E-- -CKICM---- -V-F-PCGH- --C--CA--- --CP-C
```

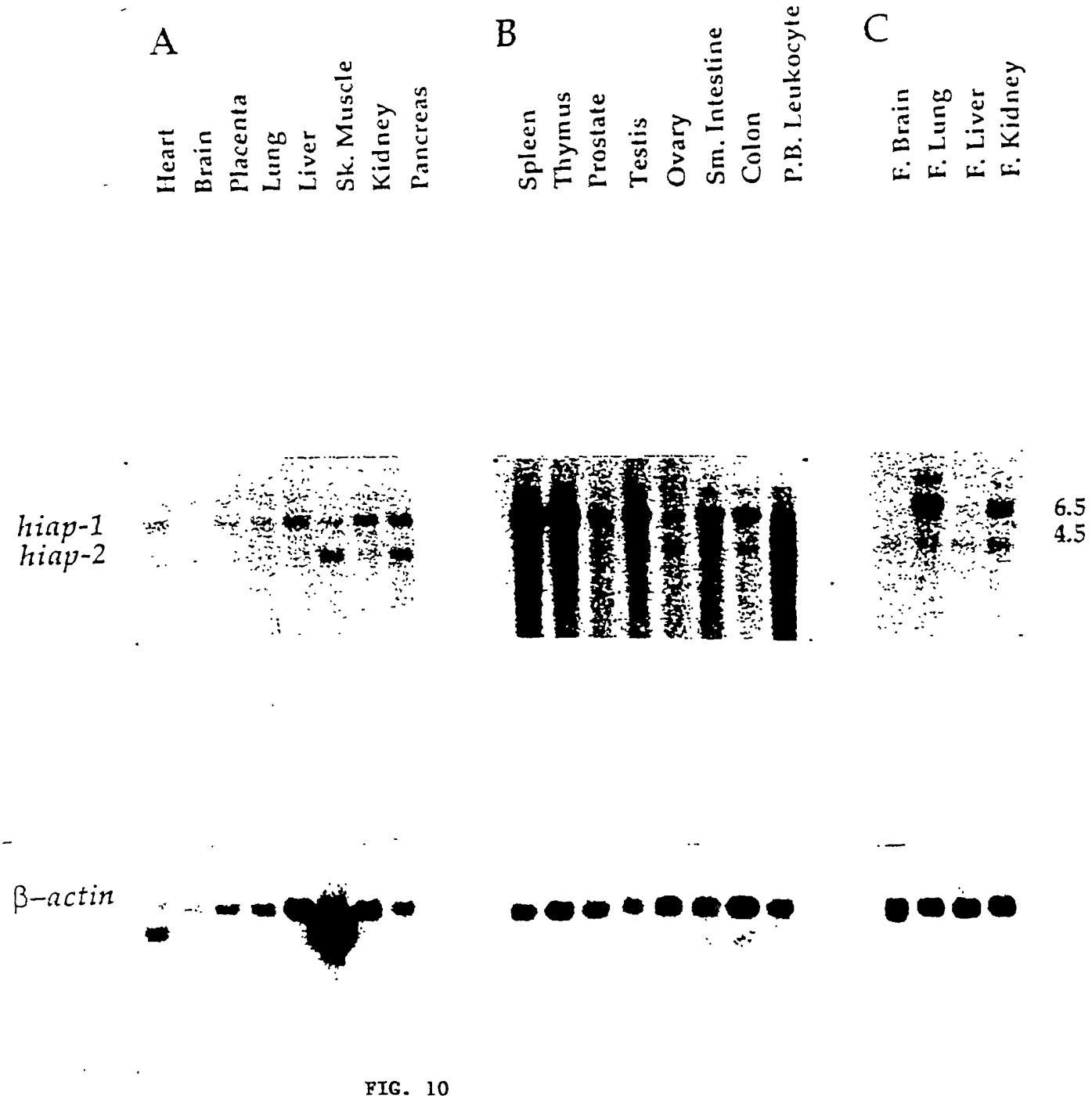
A
Heart
Brain
Placenta
Lung
Liver
Sk. Muscle
Kidney
Pancreas
B
Spleen
Thymus
Prostate
Testis
Ovary
Sm. Intestine
Colon
P.B. Leukocyte
C
F. Brain
F. Lung
F. Liver
F. Kidney
hiap-1
hiap-2
6.5
4.5
β–actin

FIG. 10

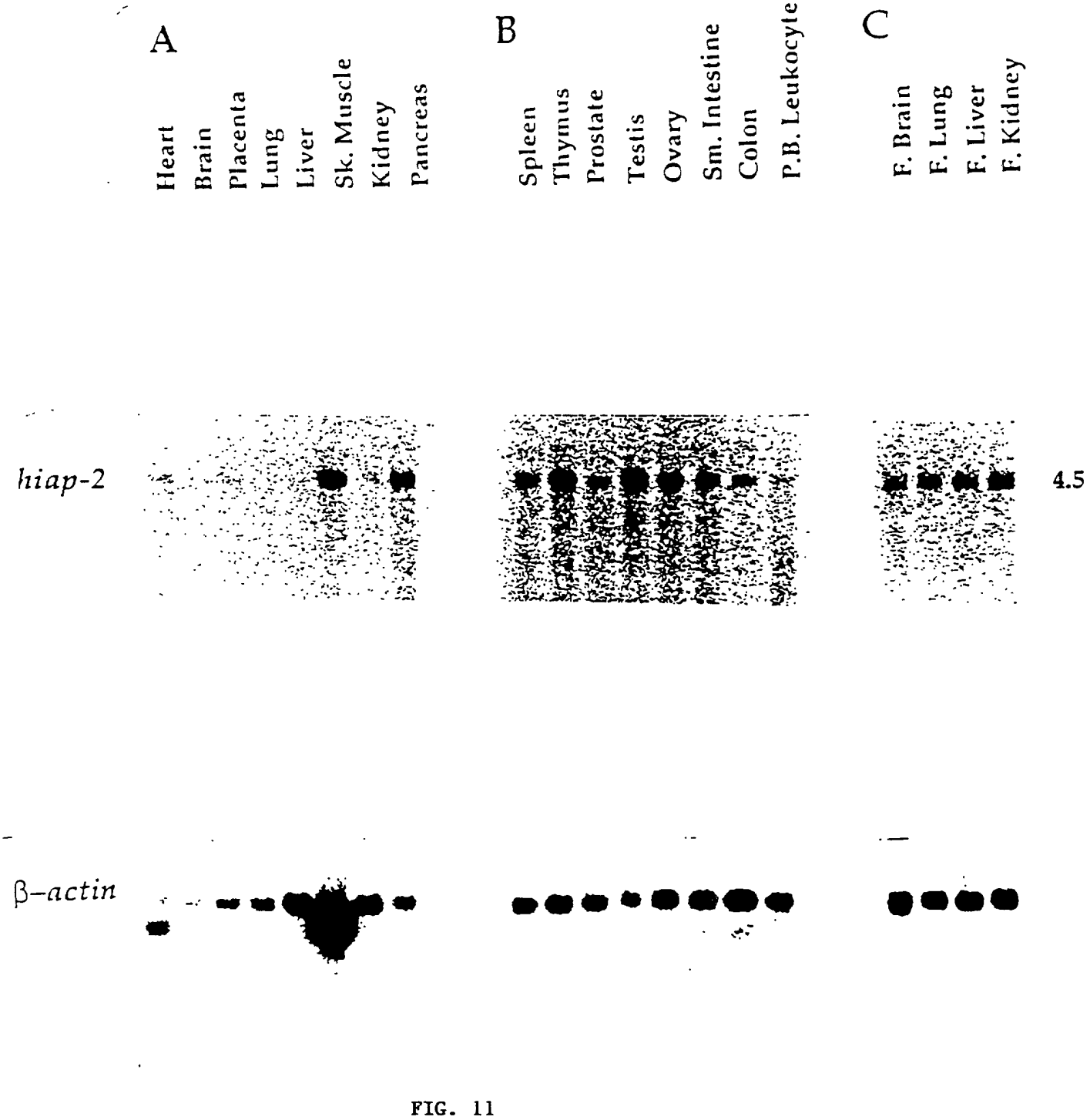
A
Heart
Brain
Placenta
Lung
Liver
Sk. Muscle
Kidney
Pancreas
B
Spleen
Thymus
Prostate
Testis
Ovary
Sm. Intestine
Colon
P.B. Leukocyte
C
F. Brain
F. Lung
F. Liver
F. Kidney
hiap-2
4.5
β–actin

FIG. 11

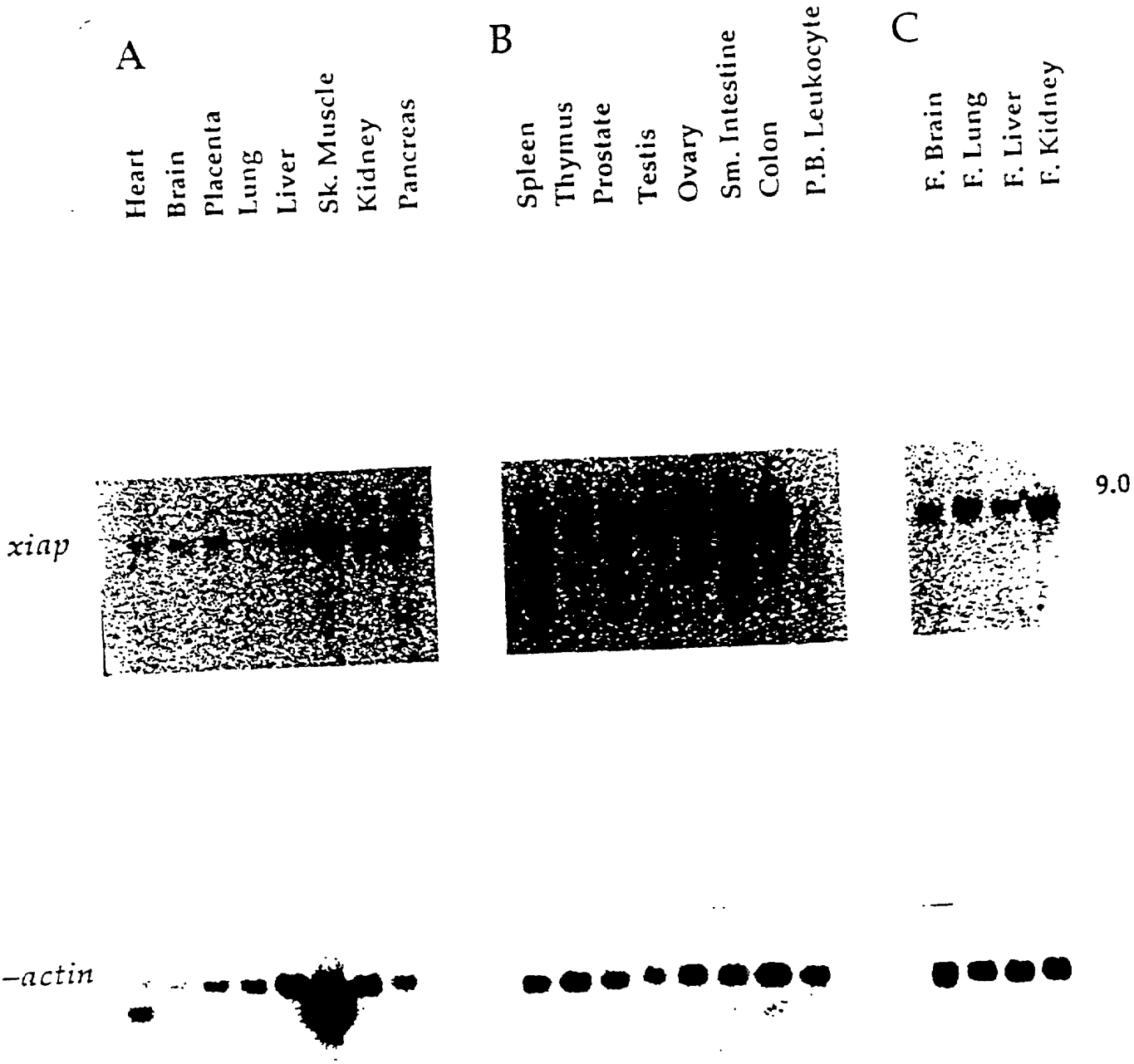
A
Heart
Brain
Placenta
Lung
Liver
Sk. Muscle
Kidney
Pancreas
B
Spleen
Thymus
Prostate
Testis
Ovary
Sm. Intestine
Colon
P.B. Leukocyte
C
F. Brain
F. Lung
F. Liver
F. Kidney
xiap
9.0
β–actin

FIG. 12

13A

13B

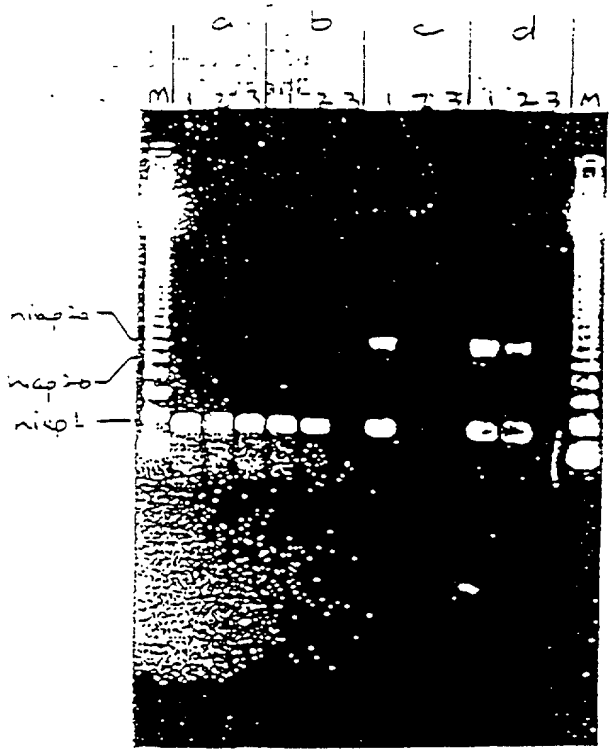

Fig. 13A and 13B

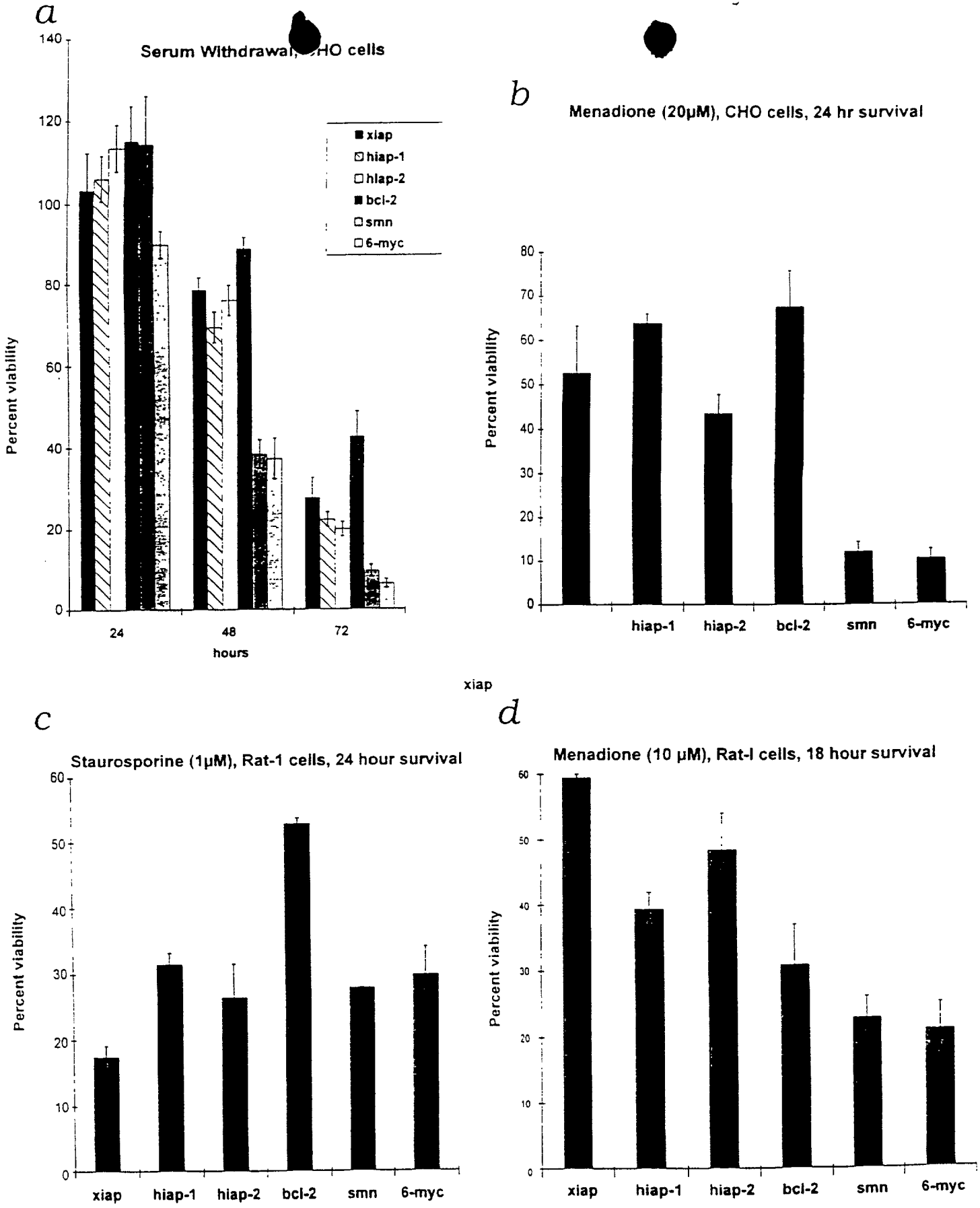
a
Serum Withdrawal, CHO cells
Percent viability
hours
24
48
72
xiap
hiap-1
hiap-2
bcl-2
smn
6-myc
b
Menadione (20μM), CHO cells, 24 hr survival
Percent viability
hiap-1
hiap-2
bcl-2
smn
6-myc
c
Staurosporine (1μM), Rat-1 cells, 24 hour survival
Percent viability
xiap
hiap-1
hiap-2
bcl-2
smn
6-myc
d
Menadione (10 μM), Rat-I cells, 18 hour survival
Percent viability
xiap
hiap-1
hiap-2
bcl-2
smn
6-myc

Fig. 14A - D

COMBINED DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled MAMMALIAN IAP GENE FAMILY, PRIMERS, PROBES AND DETECTIONS METHODS, the specification of which

☐ is attached hereto.

■ was filed on December 22, 1995 as Application Serial No. 08/576,956 and was amended on _____.

☐ was described and claimed in PCT International Application No. _____ filed on _____ and as amended under PCT Article 19 on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose all information I know to be material to patentability in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose all information I know to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56(a) which became available between the filing date of the prior application and the national or PCT international filing date of this application:

U.S. SERIAL NO.	FILING DATE	STATUS
08/511,485	August 4, 1995	■ Pending ☐ Issued ☐ Abandoned

I hereby appoint the following attorneys and/or agents to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith: Paul T. Clark, Reg. No. 30,162 and Kristina Bieker-Brady, Reg. No. 39,109, William E. Booth, Reg. No. 28,933; Barry E. Bretschneider, Reg. No. 28,055; John W. Freeman, Reg. No. 29,066; Timothy A. French, Reg. No. 30,175; Alan H. Gordon, Reg. No. 26,168; John F. Land, Reg. No. 29,554; John B. Pegram, Reg. No. 25,198; Rene D. Tegtmeyer, Reg. No. 33,567; Hans R. Troesch, Reg. No. 36,950; Dorothy P. Whelan, Reg. No. 33,814; Charles C. Winchester, Reg. No. 21,040.

Address all telephone calls to Kristina Bieker-Brady at telephone number 617/542-5070.

Address all correspondence to Kristina Bieker-Brady, Fish & Richardson P.C., 225 Franklin Street , Boston, MA 02110-2804.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patents issued thereon.

COMBINED DECL. :ATION AND POWER OF ATTOR :Y CONTINUED

Full Name of Inventor: Robert G. Korneluk

Inventor's Signature: _____ Date: Mar. 18/96

Residence Address: Ottawa, Ontario, Canada

Citizen of: Canada

Post Office Address: 1901 Tweed Ave., Ottawa, Ontario, Canada K1G 2L8

Full Name of Inventor: Alexander E. MacKenzie

Inventor's Signature: _____ Date: March 18, 96

Residence Address: Ottawa, Ontario, Canada

Citizen of: Canada

Post Office Address: 35 Rockliffe Way, Ottawa, Ontario, Canada K1M 1A3

Full Name of Inventor: Stephen Baird

Inventor's Signature: _____ Date: Mar 18/96

Residence Address: Ottawa, Ontario, Canada

Citizen of: Canada

Post Office Address: 20 Julian Ave., Ottawa, Ontario, Canada K1Y 0S5

167882.B11

SEQUENCE LISTING

<110> Korneluk, Robert G.
Mackenzie, Alexander E.
Baird, Stephen
Liston, Peter

<120> MAMMALIAN IAP GENE FAMILY, PRIMERS,
PROBES, AND DETECTION METHODS

<130> 07891/003005

<150> 08/576,956
<151> 1995-12-22

<150> 08/511,485
<151> 1995-08-04

<160> 92

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 46
<212> PRT
<213> artificial sequence based on Homo sapiens, Mus musculus, Cydia pomonella, Orgyia pseudotsugata, and Drosophila melanogaster.

<220>
<221> VARIANT
<222> 8
<223> Glu or Asp

<221> VARIANT
<222> 14,22
<223> Val or Ile

<221> VARIANT
<222> (1)...(46)
<223> Xaa = Any Amino Acid

<400> 1

```
Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Lys Xaa Cys Met
 1               5                  10                  15
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Xaa Pro Cys Gly His Xaa Xaa Xaa
            20                  25                  30
Cys Xaa Xaa Cys Ala Xaa Xaa Xaa Xaa Xaa Cys Pro Xaa Cys
        35                  40                  45
```

<210> 2
<211> 68
<212> PRT
<213> artificial sequence based on Homo sapiens, Mus musculus, Cydia pomonella, Orgyia pseudotsugata, and Drosophila melanogaster.

<220>
<221> VARIANT
<222> 13,16,17
<223> any amino acid or absent

<221> VARIANT
<222> (1)...(68)
<223> Xaa = Any Amino Acid

<400> 2
Xaa Xaa Xaa Arg Leu Xaa Thr Phe Xaa Xaa Trp Pro Xaa Xaa Xaa Xaa
 1   5   10   15
Xaa Xaa Xaa Xaa Xaa Leu Ala Xaa Ala Gly Phe Tyr Tyr Xaa Gly Xaa
            20   25   30
Xaa Asp Xaa Val Xaa Cys Phe Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Trp
        35   40   45
Xaa Xaa Xaa Asp Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Pro Xaa
    50   55   60
Cys Xaa Phe Val
65

<210> 3
<211> 2540
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(2540)
<223> n = A,T,C, or G

<400> 3
gaaaaggtgg acaagtccta ttttcaagag aagatgactt ttaacagttt tgaaggatct 60
aaaacttgtg tacctgcaga catcaataag gaagaagaat ttgtagaaga gtttaataga 120
ttaaaaactt ttgctaattt tccaagtggt agtcctgttt cagcatcaac actggcacga 180
gcagggtttc tttatactgg tgaaggagat accgtgcggt gctttagttg tcatgcagct 240
gtagatagat ggcaatatgg agactcagca gttggaagac acaggaaagt atccccaaat 300
tgcagattta tcaacggctt ttatcttgaa aatagtgcca cgcagtctac aaattctggt 360
atccagaatg gtcagtacaa agttgaaaac tatctgggaa gcagagatca ttttgcctta 420
gacaggccat ctgagacaca tgcagactat cttttgagaa ctgggcaggt tgtagatata 480
tcagacacca tatacccgag gaaccctgcc atgtattgtg aagaagctag attaaagtcc 540
tttcagaact ggccagacta tgctcaccta accccaagag agttagcaag tgctggactc 600
tactacacag gtattggtga ccaagtgcag tgcttttgtt gtggtggaaa actgaaaaat 660
tgggaacctt gtgatcgtgc ctggtcagaa cacaggcgac actttcctaa ttgcttcttt 720
gttttgggcc ggaatcttaa tattcgaagt gaatctgatg ctgtgagttc tgataggaat 780
ttcccaaatt caacaaatct tccaagaaat ccatccatgg cagattatga agcacggatc 840

tttacttttg ggacatggat atactcagtt aacaaggagc agcttgcaag agctggattt 900
tatgctttag gtgaaggtga taaagtaaag tgctttcact gtggaggagg gctaactgat 960
tggaagccca gtgaagaccc ttgggaacaa catgctaaat ggtatccagg gtgcaaatat 1020
ctgttagaac agaaggaca agaatatata aacaatattc atttaactca ttcacttgag 1080
gagtgtctgg taagaactac tgagaaaaca ccatcactaa ctagaagaat tgatgatacc 1140
atcttccaaa atcctatggt acaagaagct atacgaatgg ggttcagttt caaggacatt 1200
aagaaaataa tggaggaaaa aattcagata tctgggagca actataaatc acttgaggtt 1260
ctggttgcag atctagtgaa tgctcagaaa gacagtatgc aagatgagtc aagtcagact 1320
tcattacaga aagagattag tactgaagag cagctaaggc gcctgcaaga ggagaagctt 1380
tgcaaaatct gtatggatag aaatattgct atcgtttttg ttccttgtgg acatctagtc 1440
acttgtaaac aatgtgctga agcagttgac aagtgtccca tgtgctacac agtcattact 1500
ttcaagcaaa aaatttttat gtcttaatct aactctatag taggcatgtt atgttgttct 1560
tattaccctg attgaatgtg tgatgtgaac tgactttaag taatcaggat tgaattccat 1620
tagcatttgc taccaagtag gaaaaaaaat gtacatggca gtgttttagt tggcaatata 1680
atctttgaat ttcttgattt ttcagggtat tagctgtatt atccattttt tttactgtta 1740
tttaattgaa accatagact aagaataaga agcatcatac tataactgaa cacaatgtgt 1800
attcatagta tactgattta atttctaagt gtaagtgaat taatcatctg gatttttat 1860
tcttttcaga taggcttaac aaatggagct ttctgtatat aaatgtggag attagagtta 1920
atctccccaa tcacataatt tgttttgtgt gaaaaaggaa taaattgttc catgctggtg 1980
gaaagataga gattgttttt agaggttggt tgttgtgttt taggattctg tccattttct 2040
tgtaaaggga taaacacgga cgtgtgcgaa atatgtttgt aaagtgattt gccattgttg 2100
aaagcgtatt taatgataga atactatcga gccaacatgt actgacatgg aaagatgtca 2160
gagatatgtt aagtgtaaaa tgcaagtggc gggacactat gtatagtctg agccagatca 2220
aagtatgtat gttgttaata tgcatagaac gagagatttg gaaagatata caccaaactg 2280
ttaaatgtgg tttctcttcg gggagggggg gattggggga ggggccccag aggggtttta 2340
gaggggcctt ttcactttcg actttttca ttttgttctg ttcggatttt ttataagtat 2400
gtagaccccg aagggtttta tgggaactaa catcagtaac ctaacccccg tgactatcct 2460
gtgctcttcc tagggagctg tgttgtttcc cacccaccac ccttccctct gaacaaatgc 2520
ctgagtgctg ggcactttn 2540

<210> 4
<211> 497
<212> PRT
<213> Homo sapiens

<400> 4
Met Thr Phe Asn Ser Phe Glu Gly Ser Lys Thr Cys Val Pro Ala Asp
1 5 10 15
Ile Asn Lys Glu Glu Glu Phe Val Glu Glu Phe Asn Arg Leu Lys Thr
20 25 30
Phe Ala Asn Phe Pro Ser Gly Ser Pro Val Ser Ala Ser Thr Leu Ala
35 40 45
Arg Ala Gly Phe Leu Tyr Thr Gly Glu Gly Asp Thr Val Arg Cys Phe
50 55 60
Ser Cys His Ala Ala Val Asp Arg Trp Gln Tyr Gly Asp Ser Ala Val
65 70 75 80
Gly Arg His Arg Lys Val Ser Pro Asn Cys Arg Phe Ile Asn Gly Phe
85 90 95
Tyr Leu Glu Asn Ser Ala Thr Gln Ser Thr Asn Ser Gly Ile Gln Asn
100 105 110
Gly Gln Tyr Lys Val Glu Asn Tyr Leu Gly Ser Arg Asp His Phe Ala
115 120 125
Leu Asp Arg Pro Ser Glu Thr His Ala Asp Tyr Leu Leu Arg Thr Gly
130 135 140

```
Gln Val Val Asp Ile Ser Asp Thr Ile Tyr Pro Arg Asn Pro Ala Met
145                 150                 155                 160
Tyr Cys Glu Glu Ala Arg Leu Lys Ser Phe Gln Asn Trp Pro Asp Tyr
                165                 170                 175
Ala His Leu Thr Pro Arg Glu Leu Ala Ser Ala Gly Leu Tyr Tyr Thr
            180                 185                 190
Gly Ile Gly Asp Gln Val Gln Cys Phe Cys Cys Gly Gly Lys Leu Lys
        195                 200                 205
Asn Trp Glu Pro Cys Asp Arg Ala Trp Ser Glu His Arg Arg His Phe
    210                 215                 220
Pro Asn Cys Phe Phe Val Leu Gly Arg Asn Leu Asn Ile Arg Ser Glu
225                 230                 235                 240
Ser Asp Ala Val Ser Ser Asp Arg Asn Phe Pro Asn Ser Thr Asn Leu
                245                 250                 255
Pro Arg Asn Pro Ser Met Ala Asp Tyr Glu Ala Arg Ile Phe Thr Phe
            260                 265                 270
Gly Thr Trp Ile Tyr Ser Val Asn Lys Glu Gln Leu Ala Arg Ala Gly
        275                 280                 285
Phe Tyr Ala Leu Gly Glu Gly Asp Lys Val Lys Cys Phe His Cys Gly
    290                 295                 300
Gly Gly Leu Thr Asp Trp Lys Pro Ser Glu Asp Pro Trp Glu Gln His
305                 310                 315                 320
Ala Lys Trp Tyr Pro Gly Cys Lys Tyr Leu Leu Glu Gln Lys Gly Gln
                325                 330                 335
Glu Tyr Ile Asn Asn Ile His Leu Thr His Ser Leu Glu Glu Cys Leu
            340                 345                 350
Val Arg Thr Thr Glu Lys Thr Pro Ser Leu Thr Arg Arg Ile Asp Asp
        355                 360                 365
Thr Ile Phe Gln Asn Pro Met Val Gln Glu Ala Ile Arg Met Gly Phe
    370                 375                 380
Ser Phe Lys Asp Ile Lys Lys Ile Met Glu Glu Lys Ile Gln Ile Ser
385                 390                 395                 400
Gly Ser Asn Tyr Lys Ser Leu Glu Val Leu Val Ala Asp Leu Val Asn
                405                 410                 415
Ala Gln Lys Asp Ser Met Gln Asp Glu Ser Ser Gln Thr Ser Leu Gln
            420                 425                 430
Lys Glu Ile Ser Thr Glu Glu Gln Leu Arg Arg Leu Gln Glu Glu Lys
        435                 440                 445
Leu Cys Lys Ile Cys Met Asp Arg Asn Ile Ala Ile Val Phe Val Pro
    450                 455                 460
Cys Gly His Leu Val Thr Cys Lys Gln Cys Ala Glu Ala Val Asp Lys
465                 470                 475                 480
Cys Pro Met Cys Tyr Thr Val Ile Thr Phe Lys Gln Lys Ile Phe Met
                485                 490                 495
Ser
```

<210> 5
<211> 2676
<212> DNA
<213> Homo sapiens

<220>
<221> Variant
<222> (1)...(2676)
<223> n = A,T,C,or G

<400> 5
tccttgagat gtatcagtat aggatttagg atctccatgt tggaactcta aatgcataga 60
aatggaaata atggaaattt ttcattttgg cttttcagcc tagtattaaa actgataaaa 120
gcaaagccat gcacaaaact acctccctag agaaaggcta gtccctttc ttccccattc 180
atttcattat gaacatagta gaaaacagca tattcttatc aaatttgatg aaaagcgcca 240
acacgtttga actgaaatac gacttgtcat gtgaactgta ccgaatgtct acgtattcca 300
cttttcctgc tggggttcct gtctcagaaa ggagtcttgc tcgtgctggt ttctattaca 360
ctggtgtgaa tgacaaggtc aaatgcttct gttgtggcct gatgctggat aactggaaaa 420
gaggagacag tcctactgaa aagcataaaa agttgtatcc tagctgcaga ttcgttcaga 480
gtctaaattc cgttaacaac ttggaagcta cctctcagcc tacttttcct tcttcagtaa 540
cacattccac acactcatta cttccgggta cagaaaacag tggatatttc cgtggctctt 600
attcaaactc tccatcaaat cctgtaaact ccagagcaaa tcaagaattt tctgccttga 660
tgagaagttc ctacccctgt ccaatgaata acgaaaatgc cagattactt acttttcaga 720
catggccatt gacttttctg tcgccaacag atctggcacg agcaggcttt tactacatag 780
gacctggaga cagagtggct tgctttgcct gtggtggaaa attgagcaat tgggaaccga 840
aggataatgc tatgtcagaa cacctgagac attttcccaa atgcccattt atagaaaatc 900
agcttcaaga cacttcaaga tacacagttt ctaatctgag catgcagaca catgcagccc 960
gctttaaaac attctttaac tggccctcta gtgttctagt taatcctgag cagcttgcaa 1020
gtgcgggttt ttattatgtg ggtaacagtg atgatgtcaa atgcttttgc tgtgatggtg 1080
gactcaggtg ttgggaatct ggagatgatc catgggttca acatgccaag tggtttccaa 1140
ggtgtgagta cttgataaga attaaaggac aggagttcat ccgtcaagtt caagccagtt 1200
accctcatct acttgaacag ctgctatcca catcagacag cccaggagat gaaaatgcag 1260
agtcatcaat tatccatttg gaacctggag aagaccattc agaagatgca atcatgatga 1320
atactcctgt gattaatgct gccgtggaaa tgggctttag tagaagcctg gtaaaacaga 1380
cagttcagag aaaaatccta gcaactggag agaattatag actagtcaat gatcttgtgt 1440
tagacttact caatgcagaa gatgaaataa gggaagagga gagagaaaga gcaactgagg 1500
aaaaagaatc aaatgattta ttattaatcc ggaagaatag aatggcactt tttcaacatt 1560
tgacttgtgt aattccaatc ctggatagtc tactaactgc cggaattatt aatgaacaag 1620
aacatgatgt tattaaacag aagacacaga cgtctttaca agcaagagaa ctgattgata 1680
cgattttagt aaaaggaaat attgcagcca ctgtattcag aaactctctg caagaagctg 1740
aagctgtgtt atatgagcat ttatttgtgc aacaggacat aaaatatatt cccacagaag 1800
atgtttcaga tctaccagtg gaagaacaat tgcggagact accagaagaa agaacatgta 1860
aagtgtgtat ggacaaagaa gtgtccatag tgtttattcc ttgtggtcat ctagtagtat 1920
gcaaagattg tgctccttct ttaagaaagt gtcctatttg taggagtaca atcaagggta 1980
cagttcgtac atttctttca tgaagaagaa ccaaaacatc gtctaaactt tagaattaat 2040
ttattaaatg tattataact ttaactttta tcctaatttg gtttccttaa aatttttatt 2100
tatttacaac tcaaaaaaca ttgttttgtg taacatattt atatatgtat ctaaaccata 2160
tgaacatata ttttttagaa actaagagaa tgataggctt ttgttcttat gaacgaaaaa 2220
gaggtagcac tacaaacaca atattcaatc caaatttcag cattattgaa attgtaagtg 2280
aagtaaaact taagatattt gagttaacct ttaagaattt taaatatttt ggcattgtac 2340
taataccggg aacatgaagc caggtgtggt ggtatgtacc tgtagtccca ggctgaggca 2400
agagaattac ttgagcccag gagtttgaat ccatcctggg cagcatactg agaccctgcc 2460
tttaaaaacn aacagnacca aanccaaaca ccagggacac atttctctgt ctttttttgat 2520
cagtgtccta tacatcgaag gtgtgcatat atgttgaatc acattttagg gacatggtgt 2580
ttttataaag aattctgtga gnaaaaattt aataaagcaa ccaaattact cttaaaaaaa 2640
aaaaaaaaaa aaaaaactcg aggggcccgt accaat 2676

<210> 6
<211> 604
<212> PRT
<213> Homo sapiens

<400> 6
Met Asn Ile Val Glu Asn Ser Ile Phe Leu Ser Asn Leu Met Lys Ser
1 5 10 15
Ala Asn Thr Phe Glu Leu Lys Tyr Asp Leu Ser Cys Glu Leu Tyr Arg
20 25 30
Met Ser Thr Tyr Ser Thr Phe Pro Ala Gly Val Pro Val Ser Glu Arg
35 40 45
Ser Leu Ala Arg Ala Gly Phe Tyr Tyr Thr Gly Val Asn Asp Lys Val
50 55 60
Lys Cys Phe Cys Cys Gly Leu Met Leu Asp Asn Trp Lys Arg Gly Asp
65 70 75 80
Ser Pro Thr Glu Lys His Lys Lys Leu Tyr Pro Ser Cys Arg Phe Val
85 90 95
Gln Ser Leu Asn Ser Val Asn Asn Leu Glu Ala Thr Ser Gln Pro Thr
100 105 110
Phe Pro Ser Ser Val Thr His Ser Thr His Ser Leu Leu Pro Gly Thr
115 120 125
Glu Asn Ser Gly Tyr Phe Arg Gly Ser Tyr Ser Asn Ser Pro Ser Asn
130 135 140
Pro Val Asn Ser Arg Ala Asn Gln Glu Phe Ser Ala Leu Met Arg Ser
145 150 155 160
Ser Tyr Pro Cys Pro Met Asn Asn Glu Asn Ala Arg Leu Leu Thr Phe
165 170 175
Gln Thr Trp Pro Leu Thr Phe Leu Ser Pro Thr Asp Leu Ala Arg Ala
180 185 190
Gly Phe Tyr Tyr Ile Gly Pro Gly Asp Arg Val Ala Cys Phe Ala Cys
195 200 205
Gly Gly Lys Leu Ser Asn Trp Glu Pro Lys Asp Asn Ala Met Ser Glu
210 215 220
His Leu Arg His Phe Pro Lys Cys Pro Phe Ile Glu Asn Gln Leu Gln
225 230 235 240
Asp Thr Ser Arg Tyr Thr Val Ser Asn Leu Ser Met Gln Thr His Ala
245 250 255
Ala Arg Phe Lys Thr Phe Phe Asn Trp Pro Ser Ser Val Leu Val Asn
260 265 270
Pro Glu Gln Leu Ala Ser Ala Gly Phe Tyr Tyr Val Gly Asn Ser Asp
275 280 285
Asp Val Lys Cys Phe Cys Cys Asp Gly Gly Leu Arg Cys Trp Glu Ser
290 295 300
Gly Asp Asp Pro Trp Val Gln His Ala Lys Trp Phe Pro Arg Cys Glu
305 310 315 320
Tyr Leu Ile Arg Ile Lys Gly Gln Glu Phe Ile Arg Gln Val Gln Ala
325 330 335
Ser Tyr Pro His Leu Leu Glu Gln Leu Leu Ser Thr Ser Asp Ser Pro
340 345 350
Gly Asp Glu Asn Ala Glu Ser Ser Ile Ile His Leu Glu Pro Gly Glu
355 360 365

```
Asp His Ser Glu Asp Ala Ile Met Met Asn Thr Pro Val Ile Asn Ala
    370                 375                 380
Ala Val Glu Met Gly Phe Ser Arg Ser Leu Val Lys Gln Thr Val Gln
385                 390                 395                 400
Arg Lys Ile Leu Ala Thr Gly Glu Asn Tyr Arg Leu Val Asn Asp Leu
                405                 410                 415
Val Leu Asp Leu Leu Asn Ala Glu Asp Glu Ile Arg Glu Glu Glu Arg
            420                 425                 430
Glu Arg Ala Thr Glu Glu Lys Glu Ser Asn Asp Leu Leu Leu Ile Arg
        435                 440                 445
Lys Asn Arg Met Ala Leu Phe Gln His Leu Thr Cys Val Ile Pro Ile
    450                 455                 460
Leu Asp Ser Leu Leu Thr Ala Gly Ile Ile Asn Glu Gln Glu His Asp
465                 470                 475                 480
Val Ile Lys Gln Lys Thr Gln Thr Ser Leu Gln Ala Arg Glu Leu Ile
                485                 490                 495
Asp Thr Ile Leu Val Lys Gly Asn Ile Ala Ala Thr Val Phe Arg Asn
            500                 505                 510
Ser Leu Gln Glu Ala Glu Ala Val Leu Tyr Glu His Leu Phe Val Gln
        515                 520                 525
Gln Asp Ile Lys Tyr Ile Pro Thr Glu Asp Val Ser Asp Leu Pro Val
    530                 535                 540
Glu Glu Gln Leu Arg Arg Leu Pro Glu Glu Arg Thr Cys Lys Val Cys
545                 550                 555                 560
Met Asp Lys Glu Val Ser Ile Val Phe Ile Pro Cys Gly His Leu Val
                565                 570                 575
Val Cys Lys Asp Cys Ala Pro Ser Leu Arg Lys Cys Pro Ile Cys Arg
            580                 585                 590
Ser Thr Ile Lys Gly Thr Val Arg Thr Phe Leu Ser
        595                 600
```

<210> 7
<211> 2580
<212> DNA
<213> Homo sapiens

<220>
<221> Variant
<222> (1)...(2580)
<223> n = A,T,C,or G

<400> 7
ttaggttacc tgaaagagtt actacaaccc caaagagttg tgttctaagt agtatcttgg 60
taattcagag agatactcat cctacctgaa tataaactga gataaatcca gtaaagaaag 120
tgtagtaaat tctacataag agtctatcat tgatttcttt ttgtggtgga aatcttagtt 180
catgtgaaga aatttcatgt gaatgtttta gctatcaaac agtactgtca cctactcatg 240
cacaaaactg cctcccaaag acttttccca ggtccctcgt atcaaaacat taagagtata 300
atggaagata gcacgatctt gtcagattgg acaaacagca acaaacaaaa aatgaagtat 360
gacttttcct gtgaactcta cagaatgtct acatattcaa ctttccccgc cggggtgcct 420
gtctcagaaa ggagtcttgc tcgtgctggt ttttattata ctggtgtgaa tgacaaggtc 480
aaatgcttct gttgtggcct gatgctggat aactggaaac taggagacag tcctattcaa 540
aagcataaac agctatatcc tagctgtagc tttattcaga atctggtttc agctagtctg 600
ggatccacct ctaagaatac gtctccaatg agaaacagtt ttgcacattc attatctccc 660

accttggaac atagtagctt gttcagtggt tcttactcca gccttcctcc aaaccctctt 720
aattctagag cagttgaaga catctcttca tcgaggacta acccctacag ttatgcaatg 780
agtactgaag aagccagatt tcttacctac catatgtggc cattaacttt tttgtcacca 840
tcagaattgg caagagctgg tttttattat ataggacctg gagatagggt agcctgcttt 900
gcctgtggtg ggaagctcag taactgggaa ccaaaggatg atgctatgtc agaacaccgg 960
aggcattttc ccaactgtcc atttttggaa aattctctag aaactctgag gtttagcatt 1020
tcaaatctga gcatgcagac acatgcagct cgaatgagaa catttatgta ctggccatct 1080
agtgttccag ttcagcctga gcagcttgca agtgctggtt tttattatgt gggtcgcaat 1140
gatgatgtca aatgctttgg ttgtgatggt ggcttgaggt gttgggaatc tggagatgat 1200
ccatgggtag aacatgccaa gtggtttcca aggtgtgagt tcttgatacg aatgaaaggc 1260
caagagtttg ttgatgagat tcaaggtaga tatcctcatc ttcttgaaca gctgttgtca 1320
acttcagata ccactggaga agaaaatgct gacccaccaa ttattcattt tggacctgga 1380
gaaagttctt cagaagatgc tgtcatgatg aatacacctg tggttaaatc tgccttggaa 1440
atgggcttta atagagacct ggtgaaacaa acagttctaa gtaaaatcct gacaactgga 1500
gagaactata aaacagttaa tgatattgtg tcagcacttc ttaatgctga agatgaaaaa 1560
agagaagagg agaaggaaaa acaagctgaa gaaatggcat cagatgattt gtcattaatt 1620
cggaagaaca gaatggctct ctttcaacaa ttgacatgtg tgcttcctat cctggataat 1680
cttttaaagg ccaatgtaat taataaacag gaacatgata ttattaaaca aaaaacacag 1740
ataccttac aagcgagaga actgattgat accatttggg ttaaaggaaa tgctgcggcc 1800
aacatcttca aaaactgtct aaaagaaatt gactctacat tgtataagaa cttatttgtg 1860
gataagaata tgaagtatat tccaacagaa gatgtttcag gtctgtcact ggaagaacaa 1920
ttgaggaggt tgcaagaaga acgaacttgt aaagtgtgta tggacaaaga agtttctgtt 1980
gtatttattc cttgtggtca tctggtagta tgccaggaat gtgccccttc tctaagaaaa 2040
tgccctattt gcaggggtat aatcaagggt actgttcgta catttctctc ttaaagaaaa 2100
atagtctata ttttaacctg cataaaaagg tctttaaaat attgttgaac acttgaagcc 2160
atctaaagta aaaagggaat tatgagtttt tcaattagta acattcatgt tctagtctgc 2220
tttggtacta ataatcttgt ttctgaaaag atggtatcat atatttaatc ttaatctgtt 2280
tatttacaag ggaagattta tgtttggtga actatattag tatgtatgtg tacctaaggg 2340
agtagcgtcn ctgcttgtta tgcatcattt caggagttac tggatttgtt gttctttcag 2400
aaagctttga anactaaatt atagtgtaga aaagaactgg aaaccaggaa ctctggagtt 2460
catcagagtt atggtgccga attgtctttg gtgcttttca cttgtgtttt aaaataagga 2520
tttttctctt atttctcccc ctagtttgtg agaaacatct caataaagtg ctttaaaaag 2580

<210> 8
<211> 618
<212> PRT
<213> Homo sapiens

<400> 8
Met His Lys Thr Ala Ser Gln Arg Leu Phe Pro Gly Pro Ser Tyr Gln
 1   5   10   15
Asn Ile Lys Ser Ile Met Glu Asp Ser Thr Ile Leu Ser Asp Trp Thr
            20   25   30
Asn Ser Asn Lys Gln Lys Met Lys Tyr Asp Phe Ser Cys Glu Leu Tyr
        35   40   45
Arg Met Ser Thr Tyr Ser Thr Phe Pro Ala Gly Val Pro Val Ser Glu
    50   55   60
Arg Ser Leu Ala Arg Ala Gly Phe Tyr Tyr Thr Gly Val Asn Asp Lys
65   70   75   80
Val Lys Cys Phe Cys Cys Gly Leu Met Leu Asp Asn Trp Lys Leu Gly
                85   90   95
Asp Ser Pro Ile Gln Lys His Lys Gln Leu Tyr Pro Ser Cys Ser Phe
            100   105   110

```
Ile Gln Asn Leu Val Ser Ala Ser Leu Gly Ser Thr Ser Lys Asn Thr
        115                 120                 125
Ser Pro Met Arg Asn Ser Phe Ala His Ser Leu Ser Pro Thr Leu Glu
    130                 135                 140
His Ser Ser Leu Phe Ser Gly Ser Tyr Ser Ser Leu Pro Pro Asn Pro
145                 150                 155                 160
Leu Asn Ser Arg Ala Val Glu Asp Ile Ser Ser Ser Arg Thr Asn Pro
                165                 170                 175
Tyr Ser Tyr Ala Met Ser Thr Glu Glu Ala Arg Phe Leu Thr Tyr His
            180                 185                 190
Met Trp Pro Leu Thr Phe Leu Ser Pro Ser Glu Leu Ala Arg Ala Gly
        195                 200                 205
Phe Tyr Tyr Ile Gly Pro Gly Asp Arg Val Ala Cys Phe Ala Cys Gly
    210                 215                 220
Gly Lys Leu Ser Asn Trp Glu Pro Lys Asp Asp Ala Met Ser Glu His
225                 230                 235                 240
Arg Arg His Phe Pro Asn Cys Pro Phe Leu Glu Asn Ser Leu Glu Thr
                245                 250                 255
Leu Arg Phe Ser Ile Ser Asn Leu Ser Met Gln Thr His Ala Ala Arg
            260                 265                 270
Met Arg Thr Phe Met Tyr Trp Pro Ser Ser Val Pro Val Gln Pro Glu
        275                 280                 285
Gln Leu Ala Ser Ala Gly Phe Tyr Tyr Val Gly Arg Asn Asp Asp Val
    290                 295                 300
Lys Cys Phe Gly Cys Asp Gly Gly Leu Arg Cys Trp Glu Ser Gly Asp
305                 310                 315                 320
Asp Pro Trp Val Glu His Ala Lys Trp Phe Pro Arg Cys Glu Phe Leu
                325                 330                 335
Ile Arg Met Lys Gly Gln Glu Phe Val Asp Glu Ile Gln Gly Arg Tyr
            340                 345                 350
Pro His Leu Leu Glu Gln Leu Leu Ser Thr Ser Asp Thr Thr Gly Glu
        355                 360                 365
Glu Asn Ala Asp Pro Pro Ile Ile His Phe Gly Pro Gly Glu Ser Ser
    370                 375                 380
Ser Glu Asp Ala Val Met Met Asn Thr Pro Val Val Lys Ser Ala Leu
385                 390                 395                 400
Glu Met Gly Phe Asn Arg Asp Leu Val Lys Gln Thr Val Leu Ser Lys
                405                 410                 415
Ile Leu Thr Thr Gly Glu Asn Tyr Lys Thr Val Asn Asp Ile Val Ser
            420                 425                 430
Ala Leu Leu Asn Ala Glu Asp Glu Lys Arg Glu Glu Glu Lys Glu Lys
        435                 440                 445
Gln Ala Glu Glu Met Ala Ser Asp Asp Leu Ser Leu Ile Arg Lys Asn
    450                 455                 460
Arg Met Ala Leu Phe Gln Gln Leu Thr Cys Val Leu Pro Ile Leu Asp
465                 470                 475                 480
Asn Leu Leu Lys Ala Asn Val Ile Asn Lys Gln Glu His Asp Ile Ile
                485                 490                 495
Lys Gln Lys Thr Gln Ile Pro Leu Gln Ala Arg Glu Leu Ile Asp Thr
            500                 505                 510
Ile Trp Val Lys Gly Asn Ala Ala Ala Asn Ile Phe Lys Asn Cys Leu
        515                 520                 525
Lys Glu Ile Asp Ser Thr Leu Tyr Lys Asn Leu Phe Val Asp Lys Asn
    530                 535                 540
```

```
Met Lys Tyr Ile Pro Thr Glu Asp Val Ser Gly Leu Ser Leu Glu Glu
545                 550                 555                 560
Gln Leu Arg Arg Leu Gln Glu Glu Arg Thr Cys Lys Val Cys Met Asp
                565                 570                 575
Lys Glu Val Ser Val Val Phe Ile Pro Cys Gly His Leu Val Val Cys
            580                 585                 590
Gln Glu Cys Ala Pro Ser Leu Arg Lys Cys Pro Ile Cys Arg Gly Ile
        595                 600                 605
Ile Lys Gly Thr Val Arg Thr Phe Leu Ser
    610                 615
```

<210> 9
<211> 2100
<212> DNA
<213> Mus musculus

<400> 9

```
gacactctgc tgggcggcgg gccgccctcc tccgggacct cccctcggga accgtcgccc 60
gcggcgctta gttaggactg gagtgcttgg cgcgaaaagg tggacaagtc ctattttcca 120
gagaagatga cttttaacag ttttgaagga actagaactt ttgtacttgc agacaccaat 180
aaggatgaag aatttgtaga agagtttaat agattaaaaa catttgctaa cttcccaagt 240
agtagtcctg tttcagcatc aacattggcg cgagctgggt ttctttatac cggtgaagga 300
gacaccgtgc aatgtttcag ttgtcatgcg gcaatagata gatggcagta tggagactca 360
gctgttggaa gacacaggag aatatcccca aattgcagat ttatcaatgg ttttttatttt 420
gaaaatggtg ctgcacagtc tacaaatcct ggtatccaaa atggccagta caaatctgaa 480
aactgtgtgg gaaatagaaa tccttttgcc cctgacaggc cacctgagac tcatgctgat 540
tatctcttga gaactggaca ggttgtagat atttcagaca ccatataccc gaggaaccct 600
gccatgtgta gtgaagaagc cagattgaag tcatttcaga actggccgga ctatgctcat 660
ttaaccccca gagagttagc tagtgctggc ctctactaca caggggctga tgatcaagtg 720
caatgctttt gttgtggggg aaaactgaaa aattgggaac cctgtgatcg tgcctggtca 780
gaacacagga gacactttcc caattgcttt tttgttttgg gccggaacgt taatgttcga 840
agtgaatctg gtgtgagttc tgataggaat ttcccaaatt caacaaactc tccaagaaat 900
ccagccatgg cagaatatga agcacggatc gttacttttg gaacatggat atactcagtt 960
aacaaggagc agcttgcaag agctggattt tatgctttag gtgaaggcga taaagtgaag 1020
tgcttccact gtggaggagg gctcacggat tggaagccaa gtgaagaccc ctgggaccag 1080
catgctaagt gctacccagg gtgcaaatac ctattggatg agaaggggca agaatatata 1140
aataatattc atttaaccca tccacttgag gaatctttgg gaagaactgc tgaaaaaaca 1200
ccaccgctaa ctaaaaaaat cgatgatacc atcttccaga atcctatggt gcaagaagct 1260
atacgaatgg gatttagctt caaggacctt aagaaaacaa tggaagaaaa aatccaaaca 1320
tccgggagca gctatctatc acttgaggtc ctgattgcag atcttgtgag tgctcagaaa 1380
gataatacgg aggatgagtc aagtcaaact tcattgcaga aagacattag tactgaagag 1440
cagctaaggc gcctacaaga ggagaagctt tccaaaatct gtatggatag aaatattgct 1500
atcgtttttt ttccttgtgg acatctggcc acttgtaaac agtgtgcaga agcagttgac 1560
aaatgtccca tgtgctacac cgtcattacg ttcaaccaaa aaatttttat gtcttagtgg 1620
ggcaccacat gttatgttct tcttgctcta attgaatgtg taatgggagc gaactttaag 1680
taatcctgca tttgcattcc attagcatcc tgctgtttcc aaatggagac caatgctaac 1740
agcactgttt ccgtctaaac attcaatttc tggatctttc gagttatcag ctgtatcatt 1800
tagccagtgt tttactcgat tgaaacctta gacagagaag cattttatag cttttcacat 1860
gtatattggt agtacactga cttgatttct atatgtaagt gaattcatca cctgcatgtt 1920
tcatgccttt tgcataagct taacaaatgg agtgttctgt ataagcatgg agatgtgatg 1980
gaatctgccc aatgacttta attggcttat tgtaaacacg gaaagaactg ccccacgctg 2040
ctgggaggat aaagattgtt ttagatgctc acttctgtgt tttaggattc tgcccattta 2100
```

<210> 10
<211> 496
<212> PRT
<213> Homo sapiens

<400> 10
Met Thr Phe Asn Ser Phe Glu Gly Thr Arg Thr Phe Val Leu Ala Asp
 1   5   10   15
Thr Asn Lys Asp Glu Glu Phe Val Glu Glu Phe Asn Arg Leu Lys Thr
            20   25   30
Phe Ala Asn Phe Pro Ser Ser Ser Pro Val Ser Ala Ser Thr Leu Ala
        35   40   45
Arg Ala Gly Phe Leu Tyr Thr Gly Glu Gly Asp Thr Val Gln Cys Phe
    50   55   60
Ser Cys His Ala Ala Ile Asp Arg Trp Gln Tyr Gly Asp Ser Ala Val
65   70   75   80
Gly Arg His Arg Arg Ile Ser Pro Asn Cys Arg Phe Ile Asn Gly Phe
                85   90   95
Tyr Phe Glu Asn Gly Ala Ala Gln Ser Thr Asn Pro Gly Ile Gln Asn
            100   105   110
Gly Gln Tyr Lys Ser Glu Asn Cys Val Gly Asn Arg Asn Pro Phe Ala
        115   120   125
Pro Asp Arg Pro Pro Glu Thr His Ala Asp Tyr Leu Leu Arg Thr Gly
    130   135   140
Gln Val Val Asp Ile Ser Asp Thr Ile Tyr Pro Arg Asn Pro Ala Met
145   150   155   160
Cys Ser Glu Glu Ala Arg Leu Lys Ser Phe Gln Asn Trp Pro Asp Tyr
                165   170   175
Ala His Leu Thr Pro Arg Glu Leu Ala Ser Ala Gly Leu Tyr Tyr Thr
            180   185   190
Gly Ala Asp Asp Gln Val Gln Cys Phe Cys Cys Gly Gly Lys Leu Lys
        195   200   205
Asn Trp Glu Pro Cys Asp Arg Ala Trp Ser Glu His Arg Arg His Phe
    210   215   220
Pro Asn Cys Phe Phe Val Leu Gly Arg Asn Val Asn Val Arg Ser Glu
225   230   235   240
Ser Gly Val Ser Ser Asp Arg Asn Phe Pro Asn Ser Thr Asn Ser Pro
                245   250   255
Arg Asn Pro Ala Met Ala Glu Tyr Glu Ala Arg Ile Val Thr Phe Gly
            260   265   270
Thr Trp Ile Tyr Ser Val Asn Lys Glu Gln Leu Ala Arg Ala Gly Phe
        275   280   285
Tyr Ala Leu Gly Glu Gly Asp Lys Val Lys Cys Phe His Cys Gly Gly
    290   295   300
Gly Leu Thr Asp Trp Lys Pro Ser Glu Asp Pro Trp Asp Gln His Ala
305   310   315   320
Lys Cys Tyr Pro Gly Cys Lys Tyr Leu Leu Asp Glu Lys Gly Gln Glu
                325   330   335
Tyr Ile Asn Asn Ile His Leu Thr His Pro Leu Glu Glu Ser Leu Gly
            340   345   350
Arg Thr Ala Glu Lys Thr Pro Pro Leu Thr Lys Lys Ile Asp Asp Thr
        355   360   365
Ile Phe Gln Asn Pro Met Val Gln Glu Ala Ile Arg Met Gly Phe Ser
    370   375   380

```
Phe Lys Asp Leu Lys Lys Thr Met Glu Glu Lys Ile Gln Thr Ser Gly
385                 390                 395                 400
Ser Ser Tyr Leu Ser Leu Glu Val Leu Ile Ala Asp Leu Val Ser Ala
                405                 410                 415
Gln Lys Asp Asn Thr Glu Asp Glu Ser Ser Gln Thr Ser Leu Gln Lys
            420                 425                 430
Asp Ile Ser Thr Glu Glu Gln Leu Arg Arg Leu Gln Glu Glu Lys Leu
        435                 440                 445
Ser Lys Ile Cys Met Asp Arg Asn Ile Ala Ile Val Phe Phe Pro Cys
    450                 455                 460
Gly His Leu Ala Thr Cys Lys Gln Cys Ala Glu Ala Val Asp Lys Cys
465                 470                 475                 480
Pro Met Cys Tyr Thr Val Ile Thr Phe Asn Gln Lys Ile Phe Met Ser
                485                 490                 495


<210> 11
<211> 67
<212> PRT
<213> Orgyia pseudotsugata

<400> 11
Lys Ala Ala Arg Leu Gly Thr Tyr Thr Asn Trp Pro Val Gln Phe Leu
 1               5                  10                  15
Glu Pro Ser Arg Met Ala Ala Ser Gly Phe Tyr Tyr Leu Gly Arg Gly
            20                  25                  30
Asp Glu Val Arg Cys Ala Phe Cys Lys Val Glu Ile Thr Asn Trp Val
        35                  40                  45
Arg Gly Asp Asp Pro Glu Thr Asp His Lys Arg Trp Ala Pro Gln Cys
    50                  55                  60
Pro Phe Val
65


<210> 12
<211> 275
<212> PRT
<213> Cydia pomonella

<400> 12
Met Ser Asp Leu Arg Leu Glu Glu Val Arg Leu Asn Thr Phe Glu Lys
 1               5                  10                  15
Trp Pro Val Ser Phe Leu Ser Pro Glu Thr Met Ala Lys Asn Gly Phe
            20                  25                  30
Tyr Tyr Leu Gly Arg Ser Asp Glu Val Arg Cys Ala Phe Cys Lys Val
        35                  40                  45
Glu Ile Met Arg Trp Lys Glu Gly Glu Asp Pro Ala Ala Asp His Lys
    50                  55                  60
Lys Trp Ala Pro Gln Cys Pro Phe Val Lys Gly Ile Asp Val Cys Gly
65                  70                  75                  80
Ser Ile Val Thr Thr Asn Asn Ile Gln Asn Thr Thr Thr His Asp Thr
                85                  90                  95
Ile Ile Gly Pro Ala His Pro Lys Tyr Ala His Glu Ala Ala Arg Val
            100                 105                 110
```

```
Lys Ser Phe His Asn Trp Pro Arg Cys Met Lys Gln Arg Pro Glu Gln
        115                 120                 125
Met Ala Asp Ala Gly Phe Phe Tyr Thr Gly Tyr Gly Asp Asn Thr Lys
    130                 135                 140
Cys Phe Tyr Cys Asp Gly Gly Leu Lys Asp Trp Glu Pro Glu Asp Val
145                 150                 155                 160
Pro Trp Glu Gln His Val Arg Trp Phe Asp Arg Cys Ala Tyr Val Gln
                165                 170                 175
Leu Val Lys Gly Arg Asp Tyr Val Gln Lys Val Ile Thr Glu Ala Cys
            180                 185                 190
Val Leu Pro Gly Glu Asn Thr Thr Val Ser Thr Ala Ala Pro Val Ser
        195                 200                 205
Glu Pro Ile Pro Glu Thr Lys Ile Glu Lys Glu Pro Gln Val Glu Asp
    210                 215                 220
Ser Lys Leu Cys Lys Ile Cys Tyr Val Glu Glu Cys Ile Val Cys Phe
225                 230                 235                 240
Val Pro Cys Gly His Val Val Ala Cys Ala Lys Cys Ala Leu Ser Val
                245                 250                 255
Asp Lys Cys Pro Met Cys Arg Lys Ile Val Thr Ser Val Leu Lys Val
            260                 265                 270
Tyr Phe Ser
        275
```

<210> 13
<211> 498
<212> PRT
<213> Drosophila melanogaster

<400> 13

```
Met Thr Glu Leu Gly Met Glu Leu Glu Ser Val Arg Leu Ala Thr Phe
 1               5                  10                  15
Gly Glu Trp Pro Leu Asn Ala Pro Val Ser Ala Glu Asp Leu Val Ala
            20                  25                  30
Asn Gly Phe Phe Ala Thr Gly Lys Trp Leu Glu Ala Glu Cys His Phe
        35                  40                  45
Cys His Val Arg Ile Asp Arg Trp Glu Tyr Gly Asp Gln Val Ala Glu
    50                  55                  60
Arg His Arg Arg Ser Ser Pro Ile Cys Ser Met Val Leu Ala Pro Asn
65                  70                  75                  80
His Cys Gly Asn Val Pro Arg Ser Gln Glu Ser Asp Asn Glu Gly Asn
                85                  90                  95
Ser Val Val Asp Ser Pro Glu Ser Cys Ser Cys Pro Asp Leu Leu Leu
            100                 105                 110
Glu Ala Asn Arg Leu Val Thr Phe Lys Asp Trp Pro Asn Pro Asn Ile
        115                 120                 125
Thr Pro Gln Ala Leu Ala Lys Ala Gly Phe Tyr Tyr Leu Asn Arg Leu
    130                 135                 140
Asp His Val Lys Cys Val Trp Cys Asn Gly Val Ile Ala Lys Trp Glu
145                 150                 155                 160
Lys Asn Asp Asn Ala Phe Glu Glu His Lys Arg Phe Phe Pro Gln Cys
                165                 170                 175
Pro Arg Val Gln Met Gly Pro Leu Ile Glu Phe Ala Thr Gly Lys Asn
            180                 185                 190
```

```
Leu Asp Glu Leu Gly Ile Gln Pro Thr Thr Leu Pro Leu Arg Pro Lys
        195                 200                 205
Tyr Ala Cys Val Asp Ala Arg Leu Arg Thr Phe Thr Asp Trp Pro Ile
    210                 215                 220
Ser Asn Ile Gln Pro Ala Ser Ala Leu Ala Gln Ala Gly Leu Tyr Tyr
225                 230                 235                 240
Gln Lys Ile Gly Asp Gln Val Arg Cys Phe His Cys Asn Ile Gly Leu
                245                 250                 255
Arg Ser Trp Gln Lys Glu Asp Glu Pro Trp Phe Glu His Ala Lys Trp
            260                 265                 270
Ser Pro Lys Cys Gln Phe Val Leu Leu Ala Lys Gly Pro Ala Tyr Val
        275                 280                 285
Ser Glu Val Leu Ala Thr Thr Ala Ala Asn Ala Ser Ser Gln Pro Ala
    290                 295                 300
Thr Ala Pro Ala Pro Thr Leu Gln Ala Asp Val Leu Met Asp Glu Ala
305                 310                 315                 320
Pro Ala Lys Glu Ala Leu Thr Leu Gly Ile Asp Gly Gly Val Val Arg
                325                 330                 335
Asn Ala Ile Gln Arg Lys Leu Leu Ser Ser Gly Cys Ala Phe Ser Thr
            340                 345                 350
Leu Asp Glu Leu Leu His Asp Ile Phe Asp Asp Ala Gly Ala Gly Ala
        355                 360                 365
Ala Leu Glu Val Arg Glu Pro Pro Glu Pro Ser Ala Pro Phe Ile Glu
    370                 375                 380
Pro Cys Gln Ala Thr Thr Ser Lys Ala Ala Ser Val Pro Ile Pro Val
385                 390                 395                 400
Ala Asp Ser Ile Pro Ala Lys Pro Gln Ala Ala Glu Ala Val Ser Asn
                405                 410                 415
Ile Ser Lys Ile Thr Asp Glu Ile Gln Lys Met Ser Val Ser Thr Pro
            420                 425                 430
Asn Gly Asn Leu Ser Leu Glu Glu Glu Asn Arg Gln Leu Lys Asp Ala
        435                 440                 445
Arg Leu Cys Lys Val Cys Leu Asp Glu Glu Val Gly Val Val Phe Leu
    450                 455                 460
Pro Cys Gly His Leu Ala Thr Cys Asn Gln Cys Ala Pro Ser Val Ala
465                 470                 475                 480
Asn Cys Pro Met Cys Arg Ala Asp Ile Lys Gly Phe Val Arg Thr Phe
                485                 490                 495
Leu Ser
```

<210> 14
<211> 67
<212> PRT
<213> Cydia pomonella

```
<400> 14
Glu Glu Val Arg Leu Asn Thr Phe Glu Lys Trp Pro Val Ser Phe Leu
 1               5                  10                  15
Ser Pro Glu Thr Met Ala Lys Asn Gly Phe Tyr Tyr Leu Gly Arg Ser
            20                  25                  30
Asp Glu Val Arg Cys Ala Phe Cys Lys Val Glu Ile Met Arg Trp Lys
        35                  40                  45
```

Glu Gly Glu Asp Pro Ala Ala Asp His Lys Lys Trp Ala Pro Gln Cys
    50   55   60
Pro Phe Val
65

<210> 15
<211> 67
<212> PRT
<213> Homo sapiens

<400> 15
Glu Ala Asn Arg Leu Val Thr Phe Lys Asp Trp Pro Asn Pro Asn Ile
 1   5   10   15
Thr Pro Gln Ala Leu Ala Lys Ala Gly Phe Tyr Tyr Leu Asn Arg Leu
            20   25   30
Asp His Val Lys Cys Val Trp Cys Asn Gly Val Ile Ala Lys Trp Glu
        35   40   45
Lys Asn Asp Asn Ala Phe Glu Glu His Lys Arg Phe Phe Pro Gln Cys
    50   55   60
Pro Arg Val
65

<210> 16
<211> 68
<212> PRT
<213> Mus musculus

<400> 16
Glu Phe Asn Arg Leu Lys Thr Phe Ala Asn Phe Pro Ser Ser Ser Pro
 1   5   10   15
Val Ser Ala Ser Thr Leu Ala Arg Ala Gly Phe Leu Tyr Thr Gly Glu
            20   25   30
Gly Asp Thr Val Gln Cys Phe Ser Cys His Ala Ala Ile Asp Arg Trp
        35   40   45
Gln Tyr Gly Asp Ser Ala Val Gly Arg His Arg Arg Ile Ser Pro Asn
    50   55   60
Cys Arg Phe Ile
65

<210> 17
<211> 68
<212> PRT
<213> Homo sapiens

<400> 17
Glu Phe Asn Arg Leu Lys Thr Phe Ala Asn Phe Pro Ser Gly Ser Pro
 1   5   10   15
Val Ser Ala Ser Thr Leu Ala Arg Ala Gly Phe Leu Tyr Thr Gly Glu
            20   25   30
Gly Asp Thr Val Arg Cys Phe Ser Cys His Ala Ala Val Asp Arg Trp
        35   40   45

Gln Tyr Gly Asp Ser Ala Val Gly Arg His Arg Lys Val Ser Pro Asn
50 55 60
Cys Arg Phe Ile
65

<210> 18
<211> 68
<212> PRT
<213> Homo sapiens

<400> 18
Glu Leu Tyr Arg Met Ser Thr Tyr Ser Thr Phe Pro Ala Gly Val Pro
1 5 10 15
Val Ser Glu Arg Ser Leu Ala Arg Ala Gly Phe Tyr Tyr Thr Gly Val
20 25 30
Asn Asp Lys Val Lys Cys Phe Cys Cys Gly Leu Met Leu Asp Asn Trp
35 40 45
Lys Arg Gly Asp Ser Pro Thr Glu Lys His Lys Lys Leu Tyr Pro Ser
50 55 60
Cys Arg Phe Val
65

<210> 19
<211> 68
<212> PRT
<213> Homo sapiens

<400> 19
Glu Leu Tyr Arg Met Ser Thr Tyr Ser Thr Phe Pro Ala Gly Val Pro
1 5 10 15
Val Ser Glu Arg Ser Leu Ala Arg Ala Gly Phe Tyr Tyr Thr Gly Val
20 25 30
Asn Asp Lys Val Lys Cys Phe Cys Cys Gly Leu Met Leu Asp Asn Trp
35 40 45
Lys Leu Gly Asp Ser Pro Ile Gln Lys His Lys Gln Leu Tyr Pro Ser
50 55 60
Cys Ser Phe Ile
65

<210> 20
<211> 68
<212> PRT
<213> Mus musculus

<400> 20
Glu Glu Ala Arg Leu Lys Ser Phe Gln Asn Trp Pro Asp Tyr Ala His
1 5 10 15
Leu Thr Pro Arg Glu Leu Ala Ser Ala Gly Leu Tyr Tyr Thr Gly Ala
20 25 30
Asp Asp Gln Val Gln Cys Phe Cys Cys Gly Gly Lys Leu Lys Asn Trp
35 40 45

Glu Pro Cys Asp Arg Ala Trp Ser Glu His Arg Arg His Phe Pro Asn
50 55 60
Cys Phe Phe Val
65

<210> 21
<211> 68
<212> PRT
<213> Homo sapiens

<400> 21
Glu Glu Ala Arg Leu Lys Ser Phe Gln Asn Trp Pro Asp Tyr Ala His
1 5 10 15
Leu Thr Pro Arg Glu Leu Ala Ser Ala Gly Leu Tyr Tyr Thr Gly Ile
20 25 30
Gly Asp Gln Val Gln Cys Phe Cys Cys Gly Gly Lys Leu Lys Asn Trp
35 40 45
Glu Pro Cys Asp Arg Ala Trp Ser Glu His Arg Arg His Phe Pro Asn
50 55 60
Cys Phe Phe Val
65

<210> 22
<211> 67
<212> PRT
<213> Homo sapiens

<400> 22
Glu Asn Ala Arg Leu Leu Thr Phe Gln Thr Trp Pro Leu Thr Phe Leu
1 5 10 15
Ser Pro Thr Asp Leu Ala Arg Ala Gly Phe Tyr Tyr Ile Gly Pro Gly
20 25 30
Asp Arg Val Ala Cys Phe Ala Cys Gly Gly Lys Leu Ser Asn Trp Glu
35 40 45
Pro Lys Asp Asn Ala Met Ser Glu His Leu Arg His Phe Pro Lys Cys
50 55 60
Pro Phe Ile
65

<210> 23
<211> 67
<212> PRT
<213> Homo sapiens

<400> 23
Glu Glu Ala Arg Phe Leu Thr Tyr His Met Trp Pro Leu Thr Phe Leu
1 5 10 15
Ser Pro Ser Glu Leu Ala Arg Ala Gly Phe Tyr Tyr Ile Gly Pro Gly
20 25 30
Asp Arg Val Ala Cys Phe Ala Cys Gly Gly Lys Leu Ser Asn Trp Glu
35 40 45

Pro Lys Asp Asp Ala Met Ser Glu His Arg Arg His Phe Pro Asn Cys
    50   55   60
Pro Phe Leu
65

<210> 24
<211> 66
<212> PRT
<213> Mus musculus

<400> 24
Tyr Glu Ala Arg Ile Val Thr Phe Gly Thr Trp Ile Tyr Ser Val Asn
 1   5   10   15
Lys Glu Gln Leu Ala Arg Ala Gly Phe Tyr Ala Leu Gly Glu Gly Asp
            20   25   30
Lys Val Lys Cys Phe His Cys Gly Gly Gly Leu Thr Asp Trp Lys Pro
        35   40   45
Ser Glu Asp Pro Trp Asp Gln His Ala Lys Cys Tyr Pro Gly Cys Lys
    50   55   60
Tyr Leu
65

<210> 25
<211> 66
<212> PRT
<213> Homo sapiens

<400> 25
Tyr Glu Ala Arg Ile Phe Thr Phe Gly Thr Trp Ile Tyr Ser Val Asn
 1   5   10   15
Lys Glu Gln Leu Ala Arg Ala Gly Phe Tyr Ala Leu Gly Glu Gly Asp
            20   25   30
Lys Val Lys Cys Phe His Cys Gly Gly Gly Leu Thr Asp Trp Lys Pro
        35   40   45
Ser Glu Asp Pro Trp Glu Gln His Ala Lys Trp Tyr Pro Gly Cys Lys
    50   55   60
Tyr Leu
65

<210> 26
<211> 68
<212> PRT
<213> Homo sapiens

<400> 26
His Ala Ala Arg Phe Lys Thr Phe Phe Asn Trp Pro Ser Ser Val Leu
 1   5   10   15
Val Asn Pro Glu Gln Leu Ala Ser Ala Gly Phe Tyr Tyr Val Gly Asn
            20   25   30
Ser Asp Asp Val Lys Cys Phe Cys Cys Asp Gly Gly Leu Arg Cys Trp
        35   40   45

Glu Ser Gly Asp Asp Pro Trp Val Gln His Ala Lys Trp Phe Pro Arg
50 55 60
Cys Glu Tyr Leu
65

<210> 27
<211> 68
<212> PRT
<213> Homo sapiens

<400> 27
His Ala Ala Arg Met Arg Thr Phe Met Tyr Trp Pro Ser Ser Val Pro
1 5 10 15
Val Gln Pro Glu Gln Leu Ala Ser Ala Gly Phe Tyr Tyr Val Gly Arg
20 25 30
Asn Asp Asp Val Lys Cys Phe Gly Cys Asp Gly Gly Leu Arg Cys Trp
35 40 45
Glu Ser Gly Asp Asp Pro Trp Val Glu His Ala Lys Trp Phe Pro Arg
50 55 60
Cys Glu Phe Leu
65

<210> 28
<211> 68
<212> PRT
<213> Orgyia pseudotsugata

<400> 28
Glu Ala Ala Arg Leu Arg Thr Phe Ala Glu Trp Pro Arg Gly Leu Lys
1 5 10 15
Gln Arg Pro Glu Glu Leu Ala Glu Ala Gly Phe Phe Tyr Thr Gly Gln
20 25 30
Gly Asp Lys Thr Arg Cys Phe Cys Cys Asp Gly Gly Leu Lys Asp Trp
35 40 45
Glu Pro Asp Asp Ala Pro Trp Gln Gln His Ala Arg Trp Tyr Asp Arg
50 55 60
Cys Glu Tyr Val
65

<210> 29
<211> 68
<212> PRT
<213> Cydia pomonella

<400> 29
Glu Ala Ala Arg Val Lys Ser Phe His Asn Trp Pro Arg Cys Met Lys
1 5 10 15
Gln Arg Pro Glu Gln Met Ala Asp Ala Gly Phe Phe Tyr Thr Gly Tyr
20 25 30
Gly Asp Asn Thr Lys Cys Phe Tyr Cys Asp Gly Gly Leu Lys Asp Trp
35 40 45

Glu Pro Glu Asp Val Pro Trp Glu Gln His Val Arg Trp Phe Asp Arg
50 55 60
Cys Ala Tyr Val
65

<210> 30
<211> 68
<212> PRT
<213> Drosophila melanogaster

<400> 30
Val Asp Ala Arg Leu Arg Thr Phe Thr Asp Trp Pro Ile Ser Asn Ile
1 5 10 15
Gln Pro Ala Ser Ala Leu Ala Gln Ala Gly Leu Tyr Tyr Gln Lys Ile
20 25 30
Gly Asp Gln Val Arg Cys Phe His Cys Asn Ile Gly Leu Arg Ser Trp
35 40 45
Gln Lys Glu Asp Glu Pro Trp Phe Glu His Ala Lys Trp Ser Pro Lys
50 55 60
Cys Gln Phe Val
65

<210> 31
<211> 66
<212> PRT
<213> Drosophila melanogaster

<400> 31
Glu Ser Val Arg Leu Ala Thr Phe Gly Glu Trp Pro Leu Asn Ala Pro
1 5 10 15
Val Ser Ala Glu Asp Leu Val Ala Asn Gly Phe Phe Gly Thr Trp Met
20 25 30
Glu Ala Glu Cys Asp Phe Cys His Val Arg Ile Asp Arg Trp Glu Tyr
35 40 45
Gly Asp Leu Val Ala Glu Arg His Arg Arg Ser Ser Pro Ile Cys Ser
50 55 60
Met Val
65

<210> 32
<211> 46
<212> PRT
<213> Homo sapiens

<400> 32
Glu Gln Leu Arg Arg Leu Gln Glu Glu Arg Thr Cys Lys Val Cys Met
 1   5   10   15
Asp Lys Glu Val Ser Val Val Phe Ile Pro Cys Gly His Leu Val Val
            20   25   30
Cys Gln Glu Cys Ala Pro Ser Leu Arg Lys Cys Pro Ile Cys
        35   40   45

<210> 33
<211> 46
<212> PRT
<213> Homo sapiens

<400> 33
Glu Gln Leu Arg Arg Leu Pro Glu Glu Arg Thr Cys Lys Val Cys Met
 1   5   10   15
Asp Lys Glu Val Ser Ile Val Phe Ile Pro Cys Gly His Leu Val Val
            20   25   30
Cys Lys Asp Cys Ala Pro Ser Leu Arg Lys Cys Pro Ile Cys
        35   40   45

<210> 34
<211> 46
<212> PRT
<213> Homo sapiens

<400> 34
Glu Gln Leu Arg Arg Leu Gln Glu Glu Lys Leu Ser Lys Ile Cys Met
 1   5   10   15
Asp Arg Asn Ile Ala Ile Val Phe Phe Pro Cys Gly His Leu Ala Thr
            20   25   30
Cys Lys Gln Cys Ala Glu Ala Val Asp Lys Cys Pro Met Cys
        35   40   45

<210> 35
<211> 46
<212> PRT
<213> Homo sapiens

<400> 35
Glu Gln Leu Arg Arg Leu Gln Glu Glu Lys Leu Cys Lys Ile Cys Met
 1   5   10   15
Asp Arg Asn Ile Ala Ile Val Phe Val Pro Cys Gly His Leu Val Thr
            20   25   30
Cys Lys Gln Cys Ala Glu Ala Val Asp Lys Cys Pro Met Cys
        35   40   45

<210> 36
<211> 46
<212> PRT
<213> Drosophila melanogaster

<400> 36
Glu Glu Asn Arg Gln Leu Lys Asp Ala Arg Leu Cys Lys Val Cys Leu
 1   5   10   15
Asp Glu Glu Val Gly Val Val Phe Leu Pro Cys Gly His Leu Ala Thr
            20   25   30
Cys Asn Gln Cys Ala Pro Ser Val Ala Asn Cys Pro Met Cys
        35   40   45

<210> 37
<211> 46
<212> PRT
<213> Cydia pomonella

<400> 37
Glu Lys Glu Pro Gln Val Glu Asp Ser Lys Leu Cys Lys Ile Cys Tyr
 1   5   10   15
Val Glu Glu Cys Ile Val Cys Phe Val Pro Cys Gly His Val Val Ala
            20   25   30
Cys Ala Lys Cys Ala Leu Ser Val Asp Lys Cys Pro Met Cys
        35   40   45

<210> 38
<211> 46
<212> PRT
<213> Orgyia pseudotsugata

<400> 38
Ala Val Glu Ala Glu Val Ala Asp Asp Arg Leu Cys Lys Ile Cys Leu
 1   5   10   15
Gly Ala Glu Lys Thr Val Cys Phe Val Pro Cys Gly His Val Val Ala
            20   25   30
Cys Gly Lys Cys Ala Ala Gly Val Thr Thr Cys Pro Val Cys
        35   40   45

<210> 39
<211> 2474
<212> DNA
<213> Mus musculus

<400> 39
gaattccggg agacctacac ccccggagat cagaggtcat tgctggcgtt cagagcctag 60
gaagtgggct gcggtatcag cctagcagta aaaccgacca gaagccatgc acaaaactac 120
atccccagag aaagacttgt cccttccccc ccctgtcatc tcaccatgaa catggttcaa 180
gacagcgcct ttctagccaa gctgatgaag agtgctgaca cctttgagtt gaagtatgac 240
ttttcctgtg agctgtaccg attgtccacg tattcagctt ttcccagggg agttcctgtg 300

tcagaaagga gtctggctcg tgctggcttt tactacactg gtgccaatga caaggtcaag 360
tgcttctgct gtggcctgat gctagacaac tggaaacaag gggacagtcc catggagaag 420
cacagaaagt tgtaccccag ctgcaacttt gtacagactt tgaatccagc caacagtctg 480
gaagctagtc ctcggccttc tcttccttcc acggcgatga gcaccatgcc tttgagcttt 540
gcaagttctg agaatactgg ctatttcagt ggctcttact cgagctttcc ctcagaccct 600
gtgaacttcc gagcaaatca agattgtcct gctttgagca caagtcccta ccactttgca 660
atgaacacag agaaggccag attactcacc tatgaaacat ggccattgtc ttttctgtca 720
ccagcaaagc tggccaaagc aggcttctac tacataggac ctggagatag agtggcctgc 780
tttgcgtgcg atgggaaact gagcaactgg gaacgtaagg atgatgctat gtcagagcac 840
cagaggcatt tccccagctg tccgttctta aaagacttgg gtcagtctgc ttcgagatac 900
actgtctcta acctgagcat gcagacacac gcagcccgta ttagaacatt ctctaactgg 960
ccttctagtg cactagttca ttcccaggaa cttgcaagtg cgggctttta ttatacagga 1020
cacagtgatg atgtcaagtg tttatgctgt gatggtgggc tgaggtgctg ggaatctgga 1080
gatgacccct gggtggaaca tgccaagtgg tttccaaggt gtgagtactt gctcagaatc 1140
aaaggccaag aatttgtcag ccaagttcaa gctggctatc ctcatctact tgagcagcta 1200
ttatctacgt cagactcccc agaagatgag aatgcagacg cagcaatcgt gcattttggc 1260
cctggagaaa gttcggaaga tgtcgtcatg atgagcacgc ctgtggttaa agcagccttg 1320
gaaatgggct tcagtaggag cctggtgaga cagacggttc agtggcagat cctggccact 1380
ggtgagaact acaggaccgt cagtgacctc gttataggct tactcgatgc agaagacgag 1440
atgagagagg agcagatgga gcaggcggcc gaggaggagg agtcagatga tctagcacta 1500
atccggaaga acaaaatggt gcttttccaa catttgacgt gtgtgacacc aatgctgtat 1560
tgcctcctaa gtgcaagggc catcactgaa caggagtgca atgctgtgaa acagaaacca 1620
cacaccttac aagcaagcac actgattgat actgtgttag caaaaggaaa cactgcagca 1680
acctcattca gaaactccct tcgggaaatt gaccctgcgt tatacagaga tatatttgtg 1740
caacaggaca ttaggagtct tcccacagat gacattgcag ctctaccaat ggaagaacag 1800
ttgcggcccc tcccggagga cagaatgtgt aaagtgtgta tggaccgaga ggtatccatc 1860
gtgttcattc cctgtggcca tctggtcgtg tgcaaagact gcgctccctc tctgaggaag 1920
tgtcccatct gtagagggac catcaagggc acagtgcgca catttctctc ctgaacaaga 1980
ctaatggtcc atggctgcaa cttcagccag gaggaagttc actgtcactc ccagttccat 2040
tcggaacttg aggccagcct ggatagcacg agacaccgcc aaacacacaa atataaacat 2100
gaaaaacttt tgtctgaagt caagaatgaa tgaattactt atataataat tttaattggt 2160
ttccttaaaa gtgctatttg ttcccaactc agaaaattgt tttctgtaaa catatttaca 2220
tactacctgc atctaaagta ttcatatatt catatattca gatgtcatga gagagggttt 2280
tgttcttgtt cctgaaaagc tggtttatca tctgatcagc atatactgcg caacgggcag 2340
ggctagaatc catgaaccaa gctgcaaaga tctcacgcta aataaggcgg aaagatttgg 2400
agaaacgaaa ggaaattctt tcctgtccaa tgtatactct tcagactaat gacctcttcc 2460
tatcaagcct tcta 2474

<210> 40
<211> 602
<212> PRT
<213> Mus musculus

<400> 40
Met Asn Met Val Gln Asp Ser Ala Phe Leu Ala Lys Leu Met Lys Ser
 1   5   10   15
Ala Asp Thr Phe Glu Leu Lys Tyr Asp Phe Ser Cys Glu Leu Tyr Arg
            20   25   30
Leu Ser Thr Tyr Ser Ala Phe Pro Arg Gly Val Pro Val Ser Glu Arg
        35   40   45
Ser Leu Ala Arg Ala Gly Phe Tyr Tyr Thr Gly Ala Asn Asp Lys Val
    50   55   60
Lys Cys Phe Cys Cys Gly Leu Met Leu Asp Asn Trp Lys Gln Gly Asp
65   70   75   80

Ser Pro Met Glu Lys His Arg Lys Leu Tyr Pro Ser Cys Asn Phe Val
85 90 95
Gln Thr Leu Asn Pro Ala Asn Ser Leu Glu Ala Ser Pro Arg Pro Ser
100 105 110
Leu Pro Ser Thr Ala Met Ser Thr Met Pro Leu Ser Phe Ala Ser Ser
115 120 125
Glu Asn Thr Gly Tyr Phe Ser Gly Ser Tyr Ser Ser Phe Pro Ser Asp
130 135 140
Pro Val Asn Phe Arg Ala Asn Gln Asp Cys Pro Ala Leu Ser Thr Ser
145 150 155 160
Pro Tyr His Phe Ala Met Asn Thr Glu Lys Ala Arg Leu Leu Thr Tyr
165 170 175
Glu Thr Trp Pro Leu Ser Phe Leu Ser Pro Ala Lys Leu Ala Lys Ala
180 185 190
Gly Phe Tyr Tyr Ile Gly Pro Gly Asp Arg Val Ala Cys Phe Ala Cys
195 200 205
Asp Gly Lys Leu Ser Asn Trp Glu Arg Lys Asp Asp Ala Met Ser Glu
210 215 220
His Gln Arg His Phe Pro Ser Cys Pro Phe Leu Lys Asp Leu Gly Gln
225 230 235 240
Ser Ala Ser Arg Tyr Thr Val Ser Asn Leu Ser Met Gln Thr His Ala
245 250 255
Ala Arg Ile Arg Thr Phe Ser Asn Trp Pro Ser Ser Ala Leu Val His
260 265 270
Ser Gln Glu Leu Ala Ser Ala Gly Phe Tyr Tyr Thr Gly His Ser Asp
275 280 285
Asp Val Lys Cys Leu Cys Cys Asp Gly Gly Leu Arg Cys Trp Glu Ser
290 295 300
Gly Asp Asp Pro Trp Val Glu His Ala Lys Trp Phe Pro Arg Cys Glu
305 310 315 320
Tyr Leu Leu Arg Ile Lys Gly Gln Glu Phe Val Ser Gln Val Gln Ala
325 330 335
Gly Tyr Pro His Leu Leu Glu Gln Leu Leu Ser Thr Ser Asp Ser Pro
340 345 350
Glu Asp Glu Asn Ala Asp Ala Ala Ile Val His Phe Gly Pro Gly Glu
355 360 365
Ser Ser Glu Asp Val Val Met Met Ser Thr Pro Val Val Lys Ala Ala
370 375 380
Leu Glu Met Gly Phe Ser Arg Ser Leu Val Arg Gln Thr Val Gln Trp
385 390 395 400
Gln Ile Leu Ala Thr Gly Glu Asn Tyr Arg Thr Val Ser Asp Leu Val
405 410 415
Ile Gly Leu Leu Asp Ala Glu Asp Glu Met Arg Glu Glu Gln Met Glu
420 425 430
Gln Ala Ala Glu Glu Glu Glu Ser Asp Asp Leu Ala Leu Ile Arg Lys
435 440 445
Asn Lys Met Val Leu Phe Gln His Leu Thr Cys Val Thr Pro Met Leu
450 455 460
Tyr Cys Leu Leu Ser Ala Arg Ala Ile Thr Glu Gln Glu Cys Asn Ala
465 470 475 480
Val Lys Gln Lys Pro His Thr Leu Gln Ala Ser Thr Leu Ile Asp Thr
485 490 495
Val Leu Ala Lys Gly Asn Thr Ala Ala Thr Ser Phe Arg Asn Ser Leu
500 505 510

Arg Glu Ile Asp Pro Ala Leu Tyr Arg Asp Ile Phe Val Gln Gln Asp
515 520 525
Ile Arg Ser Leu Pro Thr Asp Asp Ile Ala Ala Leu Pro Met Glu Glu
530 535 540
Gln Leu Arg Pro Leu Pro Glu Asp Arg Met Cys Lys Val Cys Met Asp
545 550 555 560
Arg Glu Val Ser Ile Val Phe Ile Pro Cys Gly His Leu Val Val Cys
565 570 575
Lys Asp Cys Ala Pro Ser Leu Arg Lys Cys Pro Ile Cys Arg Gly Thr
580 585 590
Ile Lys Gly Thr Val Arg Thr Phe Leu Ser
595 600

<210> 41
<211> 2416
<212> DNA
<213> Mus musculus

<400> 41
ctgtggtgga gatctattgt ccaagtggtg agaaacttca tctggaagtt taagcggtca 60
gaaatactat tactactcat ggacaaaact gtctcccaga gactcgccca aggtacctta 120
cacccaaaaa cttaaacgta taatggagaa gagcacaatc ttgtcaaatt ggacaaagga 180
gagcgaagaa aaaatgaagt ttgacttttc gtgtgaactc taccgaatgt ctacatattc 240
agcttttccc aggggagttc ctgtctcaga gaggagtctg gctcgtgctg gcttttatta 300
tacaggtgtg aatgacaaag tcaagtgctt ctgctgtggc ctgatgttgg ataactggaa 360
acaaggggac agtcctgttg aaaagcacag acagttctat cccagctgca gctttgtaca 420
gactctgctt tcagccagtc tgcagtctcc atctaagaat atgtctcctg tgaaaagtag 480
atttgcacat tcgtcacctc tggaacgagg tggcattcac tccaacctgt gctctagccc 540
tcttaattct agagcagtgg aagacttctc atcaaggatg gatccctgca gctatgccat 600
gagtacagaa gaggccagat ttcttactta cagtatgtgg cctttaagtt ttctgtcacc 660
agcagagctg gccagagctg gcttctatta catagggcct ggagacaggg tggcctgttt 720
tgcctgtggt gggaaactga gcaactggga accaaaggat tatgctatgt cagagcaccg 780
cagacatttt ccccactgtc catttctgga aaatacttca gaaacacaga ggtttagtat 840
atcaaatcta agtatgcaga cacactctgc tcgattgagg acatttctgt actggccacc 900
tagtgttcct gttcagcccg agcagcttgc aagtgctgga ttctattacg tggatcgcaa 960
tgatgatgtc aagtgccttt gttgtgatgg tggcttgaga tgttgggaac ctggagatga 1020
cccctggata gaacacgcca aatggtttcc aaggtgtgag ttcttgatac ggatgaaggg 1080
tcaggagttt gttgatgaga ttcaagctag atatcctcat cttcttgagc agctgttgtc 1140
cacttcagac accccaggag aagaaaatgc tgaccctaca gagacagtgg tgcattttgg 1200
ccctggagaa agttcgaaag atgtcgtcat gatgagcacg cctgtggtta aagcagcctt 1260
ggaaatgggc ttcagtagga gcctggtgag acagacggtt cagcggcaga tcctggccac 1320
tggtgagaac tacaggaccg tcaatgatat tgtctcagta cttttgaatg ctgaagatga 1380
gagaagagaa gaggagaagg aaagacagac tgaagagatg gcatcaggtg acttatcact 1440
gattcggaag aatagaatgg ccctctttca acagttgaca catgtccttc ctatcctgga 1500
taatcttctt gaggccagtg taattacaaa acaggaacat gatattatta gacagaaaac 1560
acagataccc ttacaagcaa gagagcttat tgacaccgtt ttagtcaagg gaaatgctgc 1620
agccaacatc ttcaaaaact ctctgaaggg aattgactcc acgttatatg aaaacttatt 1680
tgtggaaaag aatatgaagt atattccaac agaagacgtt tcaggcttgt cattggaaga 1740
gcagttgcgg agattacaag aagaacgaac ttgcaaagtg tgtatggaca gagaggtttc 1800
tattgtgttc attccgtgtg gtcatctagt agtctgccag gaatgtgccc cttctctaag 1860
gaagtgcccc atctgcaggg ggacaatcaa ggggactgtg cgcacatttc tctcatgagt 1920
gaagaatggt ctgaaagtat tgttggacat cagaagctgt cagaacaaag aatgaactac 1980
tgatttcagc tcttcagcag gacattctac tctctttcaa gattagtaat cttgctttat 2040

```
gaagggtagc attgtatatt taagcttagt ctgttgcaag ggaaggtcta tgctgttgag 2100
ctacaggact gtgtctgttc cagagcagga gttgggatgc ttgctgtatg tccttcagga 2160
cttcttggga tttgggaatt tggggaaagc tttggaatcc agtgatgtgg agctcagaaa 2220
tcctggaacc agtgactctg gtactcagta gatagggtac cctgtacttc ttggtgcttt 2280
tccagtctgg gaaataagga ggaatctgct gctggtaaaa atttgctgga tgtgagaaat 2340
agatgaaagt gtttcgggtg ggggcgtgca tcagtgtagt gtgtgcaggg atgtatgcag 2400
gccaaacact gtgtag                                                  2416

<210> 42
<211> 591
<212> PRT
<213> Mus musculus

<400> 42
Met Glu Lys Ser Thr Ile Leu Ser Asn Trp Thr Lys Glu Ser Glu Glu
 1               5                  10                  15
Lys Met Lys Phe Asp Phe Ser Cys Glu Leu Tyr Arg Met Ser Thr Tyr
            20                  25                  30
Ser Ala Phe Pro Arg Gly Val Pro Val Ser Glu Arg Ser Leu Ala Arg
        35                  40                  45
Ala Gly Phe Tyr Tyr Thr Gly Val Asn Asp Lys Val Lys Cys Phe Cys
    50                  55                  60
Cys Gly Leu Met Leu Asp Asn Trp Lys Gln Gly Asp Ser Pro Val Glu
65                  70                  75                  80
Lys His Arg Gln Phe Tyr Pro Ser Cys Ser Phe Val Gln Thr Leu Leu
                85                  90                  95
Ser Ala Ser Leu Gln Ser Pro Ser Lys Asn Met Ser Pro Val Lys Ser
            100                 105                 110
Arg Phe Ala His Ser Ser Pro Leu Glu Arg Gly Gly Ile His Ser Asn
        115                 120                 125
Leu Cys Ser Ser Pro Leu Asn Ser Arg Ala Val Glu Asp Phe Ser Ser
    130                 135                 140
Arg Met Asp Pro Cys Ser Tyr Ala Met Ser Thr Glu Glu Ala Arg Phe
145                 150                 155                 160
Leu Thr Tyr Ser Met Trp Pro Leu Ser Phe Leu Ser Pro Ala Glu Leu
                165                 170                 175
Ala Arg Ala Gly Phe Tyr Tyr Ile Gly Pro Gly Asp Arg Val Ala Cys
            180                 185                 190
Phe Ala Cys Gly Gly Lys Leu Ser Asn Trp Glu Pro Lys Asp Tyr Ala
        195                 200                 205
Met Ser Glu His Arg Arg His Phe Pro His Cys Pro Phe Leu Glu Asn
    210                 215                 220
Thr Ser Glu Thr Gln Arg Phe Ser Ile Ser Asn Leu Ser Met Gln Thr
225                 230                 235                 240
His Ser Ala Arg Leu Arg Thr Phe Leu Tyr Trp Pro Pro Ser Val Pro
                245                 250                 255
Val Gln Pro Glu Gln Leu Ala Ser Ala Gly Phe Tyr Tyr Val Asp Arg
            260                 265                 270
Asn Asp Asp Val Lys Cys Leu Cys Cys Asp Gly Gly Leu Arg Cys Trp
        275                 280                 285
Glu Pro Gly Asp Asp Pro Trp Ile Glu His Ala Lys Trp Phe Pro Arg
    290                 295                 300
Cys Glu Phe Leu Ile Arg Met Lys Gly Gln Glu Phe Val Asp Glu Ile
305                 310                 315                 320
```

```
Gln Ala Arg Tyr Pro His Leu Leu Glu Gln Leu Leu Ser Thr Ser Asp
                325                 330                 335
Thr Pro Gly Glu Glu Asn Ala Asp Pro Thr Glu Thr Val Val His Phe
            340                 345                 350
Gly Pro Gly Glu Ser Ser Lys Asp Val Val Met Met Ser Thr Pro Val
        355                 360                 365
Val Lys Ala Ala Leu Glu Met Gly Phe Ser Arg Ser Leu Val Arg Gln
    370                 375                 380
Thr Val Gln Arg Gln Ile Leu Ala Thr Gly Glu Asn Tyr Arg Thr Val
385                 390                 395                 400
Asn Asp Ile Val Ser Val Leu Leu Asn Ala Glu Asp Glu Arg Arg Glu
                405                 410                 415
Glu Glu Lys Glu Arg Gln Thr Glu Glu Met Ala Ser Gly Asp Leu Ser
            420                 425                 430
Leu Ile Arg Lys Asn Arg Met Ala Leu Phe Gln Gln Leu Thr His Val
        435                 440                 445
Leu Pro Ile Leu Asp Asn Leu Leu Glu Ala Ser Val Ile Thr Lys Gln
    450                 455                 460
Glu His Asp Ile Ile Arg Gln Lys Thr Gln Ile Pro Leu Gln Ala Arg
465                 470                 475                 480
Glu Leu Ile Asp Thr Val Leu Val Lys Gly Asn Ala Ala Ala Asn Ile
                485                 490                 495
Phe Lys Asn Ser Leu Lys Gly Ile Asp Ser Thr Leu Tyr Glu Asn Leu
            500                 505                 510
Phe Val Glu Lys Asn Met Lys Tyr Ile Pro Thr Glu Asp Val Ser Gly
        515                 520                 525
Leu Ser Leu Glu Glu Gln Leu Arg Arg Leu Gln Glu Glu Arg Thr Cys
    530                 535                 540
Lys Val Cys Met Asp Arg Glu Val Ser Ile Val Phe Ile Pro Cys Gly
545                 550                 555                 560
His Leu Val Val Cys Gln Glu Cys Ala Pro Ser Leu Arg Lys Cys Pro
                565                 570                 575
Ile Cys Arg Gly Thr Ile Lys Gly Thr Val Arg Thr Phe Leu Ser
            580                 585                 590
```

<210> 43
<211> 11
<212> PRT
<213> artificial sequence based on Homo sapiens

<400> 43
```
Met Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
 1               5                  10
```

<210> 44
<211> 635
<212> PRT
<213> artificial sequence based on Homo sapiens, Mus musculus, Cydia pomonella, and Drosophila melanogaster

<220>
<221> VARIANT
<222> 1,2,3,635
<223> any amino acid or may be absent

<221> VARIANT
<222> (1)...(635)
<223> Xaa = Any Amino Acid

<400> 44
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1   5   10   15
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Xaa Xaa Glu Xaa Xaa Arg
            20   25   30
Leu Xaa Thr Phe Xaa Xaa Phe Pro Xaa Xaa Xaa Pro Val Ser Xaa Xaa
        35   40   45
Xaa Leu Ala Arg Ala Gly Phe Xaa Tyr Thr Gly Xaa Xaa Asp Xaa Val
    50   55   60
Xaa Cys Phe Xaa Cys Xaa Xaa Xaa Xaa Asp Xaa Trp Xaa Xaa Gly Asp
65   70   75   80
Ser Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Pro Xaa Cys Xaa Phe Ile
                85   90   95
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
            100   105   110
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Xaa Xaa Xaa Xaa
        115   120   125
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Tyr Xaa Xaa Xaa Xaa Xaa
    130   135   140
Xaa Xaa Xaa Xaa Xaa Arg Xaa Xaa Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa
145   150   155   160
Xaa Xaa Xaa Xaa Xaa Xaa Asp Xaa Ser Asp Xaa Xaa Xaa Xaa Xaa Xaa
                165   170   175
Xaa Xaa Xaa Met Xaa Xaa Glu Glu Ala Arg Leu Xaa Thr Phe Xaa Xaa
            180   185   190
Trp Pro Xaa Xaa Xaa Xaa Leu Xaa Pro Xaa Glu Leu Ala Xaa Ala Gly
        195   200   205
Phe Tyr Tyr Xaa Gly Xaa Xaa Asp Xaa Val Xaa Cys Phe Xaa Cys Gly
    210   215   220
Gly Lys Leu Xaa Asn Trp Glu Pro Xaa Asp Xaa Ala Xaa Ser Glu His
225   230   235   240
Xaa Arg His Phe Pro Xaa Cys Pro Phe Val Xaa Xaa Xaa Xaa Xaa Xaa
                245   250   255
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Xaa Xaa
            260   265   270
Ser Xaa Xaa Xaa Pro Xaa Asn Pro Xaa Met Ala Xaa Xaa Xaa Ala Arg
        275   280   285
Xaa Xaa Thr Phe Xaa Xaa Trp Pro Xaa Ser Xaa Xaa Val Xaa Xaa Glu
    290   295   300
Gln Leu Ala Xaa Ala Gly Phe Tyr Tyr Xaa Gly Xaa Gly Asp Xaa Val
305   310   315   320
Lys Cys Phe Xaa Cys Xaa Gly Gly Leu Xaa Xaa Trp Xaa Xaa Xaa Asp
                325   330   335
Asp Pro Trp Xaa Gln His Ala Lys Trp Phe Pro Xaa Cys Xaa Tyr Leu
            340   345   350

Xaa Xaa Xaa Lys Gly Gln Glu Tyr Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
355 360 365
Xaa Xaa Leu Xaa Glu Xaa Leu Xaa Xaa Thr Xaa Xaa Xaa Xaa Xaa Xaa
370 375 380
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Xaa Xaa Xaa Xaa
385 390 395 400
Xaa Xaa Asp Xaa Val Xaa Xaa Xaa Xaa Pro Xaa Val Xaa Xaa Ala Xaa
405 410 415
Xaa Met Gly Phe Xaa Xaa Xaa Xaa Val Lys Xaa Xaa Xaa Xaa Xaa Lys
420 425 430
Ile Xaa Xaa Xaa Gly Xaa Xaa Tyr Xaa Xaa Xaa Xaa Xaa Leu Val Xaa
435 440 445
Asp Leu Xaa Xaa Ala Xaa Xaa Xaa Xaa Xaa Xaa Xaa Glu Xaa Xaa Xaa
450 455 460
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
465 470 475 480
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
485 490 495
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
500 505 510
Xaa Xaa Xaa Xaa Gln Xaa Xaa Leu Gln Xaa Xaa Xaa Xaa Xaa Xaa Xaa
515 520 525
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
530 535 540
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
545 550 555 560
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser Xaa Glu Glu
565 570 575
Gln Leu Arg Arg Leu Xaa Glu Glu Xaa Leu Cys Lys Xaa Cys Met Asp
580 585 590
Xaa Glu Val Xaa Xaa Val Phe Xaa Pro Cys Gly His Leu Val Xaa Cys
595 600 605
Xaa Xaa Cys Ala Xaa Ser Val Xaa Lys Cys Pro Met Cys Arg Xaa Xaa
610 615 620
Ile Xaa Xaa Xaa Xaa Xaa Xaa Phe Leu Ser Xaa
625 630 635

<210> 45
<211> 204
<212> DNA
<213> Homo sapiens

<400> 45
gagtttaata gattaaaaac ttttgctaat tttccaagtg gtagtcctgt ttcagcatca 60
acactggcac gagcagggtt tctttatact ggtgaaggag ataccgtgcg gtgctttagt 120
tgtcatgcag ctgtagatag atggcaatat ggagactcag cagttggaag acacaggaaa 180
gtatccccaa attgcagatt tatc 204

<210> 46
<211> 204
<212> DNA
<213> Homo sapiens

<400> 46
gaagaagcta gattaaagtc ctttcagaac tggccagact atgctcacct aaccccaaga 60
gagttagcaa gtgctggact ctactacaca ggtattggtg accaagtgca gtgcttttgt 120
tgtggtggaa aactgaaaaa ttgggaacct tgtgatcgtg cctggtcaga acacaggcga 180
cactttccta attgcttctt tgtt 204

<210> 47
<211> 198
<212> DNA
<213> Homo sapiens

<400> 47
tatgaagcac ggatctttac ttttgggaca tggatatact cagttaacaa ggagcagctt 60
gcaagagctg gattttatgc tttaggtgaa ggtgataaag taaagtgctt tcactgtgga 120
ggagggctaa ctgattggaa gcccagtgaa gacccttggg aacaacatgc taaatggtat 180
ccagggtgca aatatctg 198

<210> 48
<211> 138
<212> DNA
<213> Homo sapiens

<400> 48
gagcagctaa ggcgcctgca agaggagaag ctttgcaaaa tctgtatgga tagaaatatt 60
gctatcgttt ttgttccttg tggacatcta gtcacttgta aacaatgtgc tgaagcagtt 120
gacaagtgtc ccatgtgc 138

<210> 49
<211> 204
<212> DNA
<213> Mus musculus

<400> 49
gagtttaata gattaaaaac atttgctaac ttcccaagta gtagtcctgt ttcagcatca 60
acattggcgc gagctgggtt tctttatacc ggtgaaggag acaccgtgca atgtttcagt 120
tgtcatgcgg caatagatag atggcagtat ggagactcag ctgttggaag acacaggaga 180
atatccccaa attgcagatt tatc 204

<210> 50
<211> 204
<212> DNA
<213> Mus musculus

<400> 50
gaagaagcca gattgaagtc atttcagaac tggccggact atgctcattt aacccccaga 60
gagttagcta gtgctggcct ctactacaca ggggctgatg atcaagtgca atgcttttgt 120
tgtgggggaa aactgaaaaa ttgggaaccc tgtgatcgtg cctggtcaga acacaggaga 180
cactttccca attgcttttt tgtt 204

<210> 51
<211> 198
<212> DNA
<213> Mus musculus

<400> 51
tatgaagcac ggatcgttac ttttggaaca tggatatact cagttaacaa ggagcagctt 60
gcaagagctg gattttatgc tttaggtgaa ggcgataaag tgaagtgctt ccactgtgga 120
ggagggctca cggattggaa gccaagtgaa gacccctggg accagcatgc taagtgctac 180
ccagggtgca aataccta 198

<210> 52
<211> 138
<212> DNA
<213> Mus musculus

<400> 52
gagcagctaa ggcgcctaca agaggagaag ctttccaaaa tctgtatgga tagaaatatt 60
gctatcgttt tttttccttg tggacatctg gccacttgta aacagtgtgc agaagcagtt 120
gacaaatgtc ccatgtgc 138

<210> 53
<211> 204
<212> DNA
<213> Homo sapiens

<400> 53
gaactgtacc gaatgtctac gtattccact tttcctgctg gggttcctgt ctcagaaagg 60
agtcttgctc gtgctggttt ctattacact ggtgtgaatg acaaggtcaa atgcttctgt 120
tgtggcctga tgctggataa ctggaaaaga ggagacagtc ctactgaaaa gcataaaaag 180
ttgtatccta gctgcagatt cgtt 204

<210> 54
<211> 201
<212> DNA
<213> Homo sapiens

<400> 54
gaaaatgcca gattacttac ttttcagaca tggccattga cttttctgtc gccaacagat 60
ctggcacgag caggctttta ctacatagga cctggagaca gagtggcttg ctttgcctgt 120
ggtggaaaat tgagcaattg ggaaccgaag gataatgcta tgtcagaaca cctgagacat 180
tttcccaaat gcccatttat a 201

<210> 55
<211> 204
<212> DNA
<213> Homo sapiens

<400> 55
catgcagccc gctttaaaac attctttaac tggccctcta gtgttctagt taatcctgag 60
cagcttgcaa gtgcgggttt ttattatgtg ggtaacagtg atgatgtcaa atgcttttgc 120
tgtgatggtg gactcaggtg ttgggaatct ggagatgatc catgggttca acatgccaag 180
tggtttccaa ggtgtgagta cttg 204

<210> 56
<211> 138
<212> DNA
<213> Homo sapiens

<400> 56
gaacaattgc ggagactacc agaagaaaga acatgtaaag tgtgtatgga caaagaagtg 60
tccatagtgt ttattccttg tggtcatcta gtagtatgca aagattgtgc tccttcttta 120
agaaagtgtc ctatttgt 138

<210> 57
<211> 203
<212> DNA
<213> Mus musculus

<400> 57
agctgtaccg attgtccacg tattcagctt ttcccagggg agttcctgtg tcagaaagga 60
gtctggctcg tgctggcttt tactacactg gtgccaatga caaggtcaag tgcttctgct 120
gtggcctgat gctagacaac tggaaacaag gggacagtcc catggagaag cacagaaagt 180
tgtaccccag ctgcaacttt gta 203

<210> 58
<211> 201
<212> DNA
<213> Mus musculus

<400> 58
gagaaggcca gattactcac ctatgaaaca tggccattgt cttttctgtc accagcaaag 60
ctggccaaag caggcttcta ctacatagga cctggagata gagtggcctg ctttgcgtgc 120
gatgggaaac tgagcaactg ggaacgtaag gatgatgcta tgtcagagca ccagaggcat 180
ttccccagct gtccgttctt a 201

<210> 59
<211> 204
<212> DNA
<213> Mus musculus

<400> 59
cacgcagccc gtattagaac attctctaac tggccttcta gtgcactagt tcattcccag 60
gaacttgcaa gtgcgggctt ttattataca ggacacagtg atgatgtcaa gtgtttatgc 120
tgtgatggtg ggctgaggtg ctgggaatct ggagatgacc cctgggtgga acatgccaag 180
tggtttccaa ggtgtgagta cttg 204

<210> 60
<211> 138
<212> DNA
<213> Mus musculus

<400> 60
gaacagttgc ggcccctccc ggaggacaga atgtgtaaag tgtgtatgga ccgagaggta 60
tccatcgtgt tcattccctg tggccatctg gtcgtgtgca aagactgcgc tccctctctg 120
aggaagtgtc ccatctgt 138

<210> 61
<211> 204
<212> DNA
<213> Homo sapiens

<400> 61
gaactctaca gaatgtctac atattcaact ttccccgccg gggtgcctgt ctcagaaagg 60
agtcttgctc gtgctggttt ttattatact ggtgtgaatg acaaggtcaa atgcttctgt 120
tgtggcctga tgctggataa ctggaaacta ggagacagtc ctattcaaaa gcataaacag 180
ctatatccta gctgtagctt tatt 204

<210> 62
<211> 201
<212> DNA
<213> Homo sapiens

<400> 62
gaagaagcca gatttcttac ctaccatatg tggccattaa ctttttttgtc accatcagaa 60
ttggcaagag ctggttttta ttatatagga cctggagata gggtagcctg ctttgcctgt 120
ggtgggaagc tcagtaactg ggaaccaaag gatgatgcta tgtcagaaca ccggaggcat 180
tttcccaact gtccattttt g 201

<210> 63
<211> 204
<212> DNA
<213> Homo sapiens

<400> 63
catgcagctc gaatgagaac atttatgtac tggccatcta gtgttccagt tcagcctgag 60
cagcttgcaa gtgctggttt ttattatgtg ggtcgcaatg atgatgtcaa atgctttggt 120
tgtgatggtg gcttgaggtg ttgggaatct ggagatgatc catgggtaga acatgccaag 180
tggtttccaa ggtgtgagtt cttg 204

<210> 64
<211> 138
<212> DNA
<213> Mus musculus

<400> 64
gaacaattga ggaggttgca agaagaacga acttgtaaag tgtgtatgga caaagaagtt 60
tctgttgtat ttattccttg tggtcatctg gtagtatgcc aggaatgtgc cccttctcta 120
agaaaatgcc ctatttgc 138

<210> 65
<211> 204
<212> DNA
<213> Mus musculus

<400> 65
gaactctacc gaatgtctac atattcagct tttcccaggg gagttcctgt ctcagagagg 60
agtctggctc gtgctggctt ttattataca ggtgtgaatg acaaagtcaa gtgcttctgc 120
tgtggcctga tgttggataa ctggaaacaa ggggacagtc ctgttgaaaa gcacagacag 180
ttctatccca gctgcagctt tgta 204

<210> 66
<211> 201
<212> DNA
<213> Mus musculus

<400> 66
gaagaggcca gatttcttac ttacagtatg tggcctttaa gttttctgtc accagcagag 60
ctggccagag ctggcttcta ttacataggg cctggagaca gggtggcctg ttttgcctgt 120
ggtgggaaac tgagcaactg ggaaccaaag gattatgcta tgtcagagca ccgcagacat 180
tttccccact gtccatttct g 201

<210> 67
<211> 204
<212> DNA
<213> Mus musculus

<400> 67
cactctgctc gattgaggac atttctgtac tggccaccta gtgttcctgt tcagcccgag 60
cagcttgcaa gtgctggatt ctattacgtg gatcgcaatg atgatgtcaa gtgcctttgt 120
tgtgatggtg gcttgagatg ttgggaacct ggagatgacc cctggataga acacgccaaa 180
tggtttccaa ggtgtgagtt cttg 204

<210> 68
<211> 114
<212> DNA
<213> Mus musculus

<400> 68
gaacgaactt gcaaagtgtg tatggacaga gaggtttcta ttgtgttcat tccgtgtggt 60
catctagtag tctgccagga atgtgcccct tctctaagga agtgccccat ctgc 114

<210> 69
<211> 68
<212> PRT
<213> Homo sapiens

<400> 69
Glu Phe Asn Arg Leu Lys Thr Phe Ala Asn Phe Pro Ser Gly Ser Pro
1 5 10 15
Val Ser Ala Ser Thr Leu Ala Arg Ala Gly Phe Leu Tyr Thr Gly Glu
20 25 30
Gly Asp Thr Val Arg Cys Phe Ser Cys His Ala Ala Val Asp Arg Trp
35 40 45
Gln Tyr Gly Asp Ser Ala Val Gly Arg His Arg Lys Val Ser Pro Asn
50 55 60
Cys Arg Phe Ile
65

<210> 70
<211> 68
<212> PRT
<213> Homo sapiens

<400> 70
Glu Glu Ala Arg Leu Lys Ser Phe Gln Asn Trp Pro Asp Tyr Ala His
1 5 10 15
Leu Thr Pro Arg Glu Leu Ala Ser Ala Gly Leu Tyr Tyr Thr Gly Ile
20 25 30

Gly Asp Gln Val Gln Cys Phe Cys Cys Gly Gly Lys Leu Lys Asn Trp
35 40 45
Glu Pro Cys Asp Arg Ala Trp Ser Glu His Arg Arg His Phe Pro Asn
50 55 60
Cys Phe Phe Val
65

<210> 71
<211> 66
<212> PRT
<213> Homo sapiens

<400> 71
Tyr Glu Ala Arg Ile Phe Thr Phe Gly Thr Trp Ile Tyr Ser Val Asn
1 5 10 15
Lys Glu Gln Leu Ala Arg Ala Gly Phe Tyr Ala Leu Gly Glu Gly Asp
20 25 30
Lys Val Lys Cys Phe His Cys Gly Gly Gly Leu Thr Asp Trp Lys Pro
35 40 45
Ser Glu Asp Pro Trp Glu Gln His Ala Lys Trp Tyr Pro Gly Cys Lys
50 55 60
Tyr Leu
65

<210> 72
<211> 46
<212> PRT
213> Homo sapiens

<400> 72
Glu Gln Leu Arg Arg Leu Gln Glu Glu Lys Leu Cys Lys Ile Cys Met
1 5 10 15
Asp Arg Asn Ile Ala Ile Val Phe Val Pro Cys Gly His Leu Val Thr
20 25 30
Cys Lys Gln Cys Ala Glu Ala Val Asp Lys Cys Pro Met Cys
35 40 45

<210> 73
<211> 68
<212> PRT
<213> Mus musculus

<400> 73
Glu Phe Asn Arg Leu Lys Thr Phe Ala Asn Phe Pro Ser Ser Ser Pro
1 5 10 15
Val Ser Ala Ser Thr Leu Ala Arg Ala Gly Phe Leu Tyr Thr Gly Glu
20 25 30
Gly Asp Thr Val Gln Cys Phe Ser Cys His Ala Ala Ile Asp Arg Trp
35 40 45

```
Gln Tyr Gly Asp Ser Ala Val Gly Arg His Arg Arg Ile Ser Pro Asn
    50                  55                  60
Cys Arg Phe Ile
65


<210> 74
<211> 68
<212> PRT
<213> Mus musculus

<400> 74
Glu Glu Ala Arg Leu Lys Ser Phe Gln Asn Trp Pro Asp Tyr Ala His
 1               5                  10                  15
Leu Thr Pro Arg Glu Leu Ala Ser Ala Gly Leu Tyr Tyr Thr Gly Ala
            20                  25                  30
Asp Asp Gln Val Gln Cys Phe Cys Cys Gly Gly Lys Leu Lys Asn Trp
        35                  40                  45
Glu Pro Cys Asp Arg Ala Trp Ser Glu His Arg Arg His Phe Pro Asn
    50                  55                  60
Cys Phe Phe Val
65


<210> 75
<211> 66
<212> PRT
<213> Mus musculus

<400> 75
Tyr Glu Ala Arg Ile Val Thr Phe Gly Thr Trp Ile Tyr Ser Val Asn
 1               5                  10                  15
Lys Glu Gln Leu Ala Arg Ala Gly Phe Tyr Ala Leu Gly Glu Gly Asp
            20                  25                  30
Lys Val Lys Cys Phe His Cys Gly Gly Gly Leu Thr Asp Trp Lys Pro
        35                  40                  45
Ser Glu Asp Pro Trp Asp Gln His Ala Lys Cys Tyr Pro Gly Cys Lys
    50                  55                  60
Tyr Leu
65


<210> 76
<211> 46
<212> PRT
<213> Mus musculus

<400> 76
Glu Gln Leu Arg Arg Leu Gln Glu Glu Lys Leu Ser Lys Ile Cys Met
 1               5                  10                  15
Asp Arg Asn Ile Ala Ile Val Phe Phe Pro Cys Gly His Leu Ala Thr
            20                  25                  30
Cys Lys Gln Cys Ala Glu Ala Val Asp Lys Cys Pro Met Cys
        35                  40                  45
```

<210> 77
<211> 68
<212> PRT
<213> Homo sapiens

<400> 77
Glu Leu Tyr Arg Met Ser Thr Tyr Ser Thr Phe Pro Ala Gly Val Pro
 1   5   10   15
Val Ser Glu Arg Ser Leu Ala Arg Ala Gly Phe Tyr Tyr Thr Gly Val
            20   25   30
Asn Asp Lys Val Lys Cys Phe Cys Cys Gly Leu Met Leu Asp Asn Trp
        35   40   45
Lys Arg Gly Asp Ser Pro Thr Glu Lys His Lys Lys Leu Tyr Pro Ser
    50   55   60
Cys Arg Phe Val
65

<210> 78
<211> 67
<212> PRT
<213> Homo sapiens

<400> 78
Glu Asn Ala Arg Leu Leu Thr Phe Gln Thr Trp Pro Leu Thr Phe Leu
 1   5   10   15
Ser Pro Thr Asp Leu Ala Arg Ala Gly Phe Tyr Tyr Ile Gly Pro Gly
            20   25   30
Asp Arg Val Ala Cys Phe Ala Cys Gly Gly Lys Leu Ser Asn Trp Glu
        35   40   45
Pro Lys Asp Asn Ala Met Ser Glu His Leu Arg His Phe Pro Lys Cys
    50   55   60
Pro Phe Ile
65

<210> 79
<211> 68
<212> PRT
<213> Homo sapiens

<400> 79
His Ala Ala Arg Phe Lys Thr Phe Phe Asn Trp Pro Ser Ser Val Leu
 1   5   10   15
Val Asn Pro Glu Gln Leu Ala Ser Ala Gly Phe Tyr Tyr Val Gly Asn
            20   25   30
Ser Asp Asp Val Lys Cys Phe Cys Cys Asp Gly Gly Leu Arg Cys Trp
        35   40   45
Glu Ser Gly Asp Asp Pro Trp Val Gln His Ala Lys Trp Phe Pro Arg
    50   55   60
Cys Glu Tyr Leu
65

<210> 80
<211> 46
<212> PRT
<213> Homo sapiens

<400> 80
Glu Gln Leu Arg Arg Leu Pro Glu Glu Arg Thr Cys Lys Val Cys Met
 1   5   10   15
Asp Lys Glu Val Ser Ile Val Phe Ile Pro Cys Gly His Leu Val Val
            20   25   30
Cys Lys Asp Cys Ala Pro Ser Leu Arg Lys Cys Pro Ile Cys
        35   40   45

<210> 81
<211> 68
<212> PRT
<213> Mus musculus

<400> 81
Glu Leu Tyr Arg Leu Ser Thr Tyr Ser Ala Phe Pro Arg Gly Val Pro
 1   5   10   15
Val Ser Glu Arg Ser Leu Ala Arg Ala Gly Phe Tyr Tyr Thr Gly Ala
            20   25   30
Asn Asp Lys Val Lys Cys Phe Cys Cys Gly Leu Met Leu Asp Asn Trp
        35   40   45
Lys Gln Gly Asp Ser Pro Met Glu Lys His Arg Lys Leu Tyr Pro Ser
    50   55   60
Cys Asn Phe Val
65

<210> 82
<211> 67
<212> PRT
<213> Mus musculus

<400> 82
Glu Lys Ala Arg Leu Leu Thr Tyr Glu Thr Trp Pro Leu Ser Phe Leu
 1   5   10   15
Ser Pro Ala Lys Leu Ala Lys Ala Gly Phe Tyr Tyr Ile Gly Pro Gly
            20   25   30
Asp Arg Val Ala Cys Phe Ala Cys Asp Gly Lys Leu Ser Asn Trp Glu
        35   40   45
Arg Lys Asp Asp Ala Met Ser Glu His Gln Arg His Phe Pro Ser Cys
    50   55   60
Pro Phe Leu
65

<210> 83
<211> 68
<212> PRT
<213> Mus musculus

<400> 83
His Ala Ala Arg Ile Arg Thr Phe Ser Asn Trp Pro Ser Ser Ala Leu
 1   5   10   15
Val His Ser Gln Glu Leu Ala Ser Ala Gly Phe Tyr Tyr Thr Gly His
            20   25   30
Ser Asp Asp Val Lys Cys Leu Cys Cys Asp Gly Gly Leu Arg Cys Trp
        35   40   45
Glu Ser Gly Asp Asp Pro Trp Val Glu His Ala Lys Trp Phe Pro Arg
    50   55   60
Cys Glu Tyr Leu
65

<210> 84
<211> 46
<212> PRT
<213> Mus musculus

<400> 84
Glu Gln Leu Arg Pro Leu Pro Glu Asp Arg Met Cys Lys Val Cys Met
 1   5   10   15
Asp Arg Glu Val Ser Ile Val Phe Ile Pro Cys Gly His Leu Val Val
            20   25   30
Cys Lys Asp Cys Ala Pro Ser Leu Arg Lys Cys Pro Ile Cys
        35   40   45

<210> 85
<211> 68
<212> PRT
<213> Homo sapiens

<400> 85
Glu Leu Tyr Arg Met Ser Thr Tyr Ser Thr Phe Pro Ala Gly Val Pro
 1   5   10   15
Val Ser Glu Arg Ser Leu Ala Arg Ala Gly Phe Tyr Tyr Thr Gly Val
            20   25   30
Asn Asp Lys Val Lys Cys Phe Cys Cys Gly Leu Met Leu Asp Asn Trp
        35   40   45
Lys Leu Gly Asp Ser Pro Ile Gln Lys His Lys Gln Leu Tyr Pro Ser
    50   55   60
Cys Ser Phe Ile
65

<210> 86
<211> 67
<212> PRT
<213> Homo sapiens

<400> 86
Glu Glu Ala Arg Phe Leu Thr Tyr His Met Trp Pro Leu Thr Phe Leu
 1   5   10   15
Ser Pro Ser Glu Leu Ala Arg Ala Gly Phe Tyr Tyr Ile Gly Pro Gly
            20   25   30
Asp Arg Val Ala Cys Phe Ala Cys Gly Gly Lys Leu Ser Asn Trp Glu
        35   40   45
Pro Lys Asp Asp Ala Met Ser Glu His Arg Arg His Phe Pro Asn Cys
    50   55   60
Pro Phe Leu
65

<210> 87
<211> 68
<212> PRT
<213> Homo sapiens

<400> 87
His Ala Ala Arg Met Arg Thr Phe Met Tyr Trp Pro Ser Ser Val Pro
 1   5   10   15
Val Gln Pro Glu Gln Leu Ala Ser Ala Gly Phe Tyr Tyr Val Gly Arg
            20   25   30
Asn Asp Asp Val Lys Cys Phe Gly Cys Asp Gly Gly Leu Arg Cys Trp
        35   40   45
Glu Ser Gly Asp Asp Pro Trp Val Glu His Ala Lys Trp Phe Pro Arg
    50   55   60
Cys Glu Phe Leu
65

<210> 88
<211> 46
<212> PRT
<213> Homo sapiens

<400> 88
Glu Gln Leu Arg Arg Leu Gln Glu Glu Arg Thr Cys Lys Val Cys Met
 1   5   10   15
Asp Lys Glu Val Ser Val Val Phe Ile Pro Cys Gly His Leu Val Val
            20   25   30
Cys Gln Glu Cys Ala Pro Ser Leu Arg Lys Cys Pro Ile Cys
        35   40   45

<210> 89
<211> 68
<212> PRT
<213> Mus musculus

<400> 89
Glu Leu Tyr Arg Met Ser Thr Tyr Ser Ala Phe Pro Arg Gly Val Pro
1 5 10 15
Val Ser Glu Arg Ser Leu Ala Arg Ala Gly Phe Tyr Tyr Thr Gly Val
20 25 30
Asn Asp Lys Val Lys Cys Phe Cys Cys Gly Leu Met Leu Asp Asn Trp
35 40 45
Lys Gln Gly Asp Ser Pro Val Glu Lys His Arg Gln Phe Tyr Pro Ser
50 55 60
Cys Ser Phe Val
65

<210> 90
<211> 67
<212> PRT
<213> Mus musculus

<400> 90
Glu Glu Ala Arg Phe Leu Thr Tyr Ser Met Trp Pro Leu Ser Phe Leu
1 5 10 15
Ser Pro Ala Glu Leu Ala Arg Ala Gly Phe Tyr Tyr Ile Gly Pro Gly
20 25 30
Asp Arg Val Ala Cys Phe Ala Cys Gly Gly Lys Leu Ser Asn Trp Glu
35 40 45
Pro Lys Asp Tyr Ala Met Ser Glu His Arg Arg His Phe Pro His Cys
50 55 60
Pro Phe Leu
65

<210> 91
<211> 68
<212> PRT
<213> Mus musculus

<400> 91
His Ser Ala Arg Leu Arg Thr Phe Leu Tyr Trp Pro Pro Ser Val Pro
1 5 10 15
Val Gln Pro Glu Gln Leu Ala Ser Ala Gly Phe Tyr Tyr Val Asp Arg
20 25 30
Asn Asp Asp Val Lys Cys Leu Cys Cys Asp Gly Gly Leu Arg Cys Trp
35 40 45
Glu Pro Gly Asp Asp Pro Trp Ile Glu His Ala Lys Trp Phe Pro Arg
50 55 60
Cys Glu Phe Leu
65

<210> 92
<211> 38
<212> PRT
<213> Mus musculus

<400> 92

Glu Arg Thr Cys Lys Val Cys Met Asp Arg Glu Val Ser Ile Val Phe
 1   5   10   15

Ile Pro Cys Gly His Leu Val Val Cys Gln Glu Cys Ala Pro Ser Leu
            20   25   30

Arg Lys Cys Pro Ile Cys
        35

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Korneluk, Robert G.
Mackenzie, Alexander E.
Baird, Stephen

(ii) TITLE OF INVENTION: MAMMALIAN IAP GENE FAMILY, PRIMERS, PROBES, AND DETECTION METHODS

(iii) NUMBER OF SEQUENCES: 42

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Fish & Richardson P.C.
(B) STREET: 225 Franklin Street
(C) CITY: Boston
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02110-2804

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/511,485
(B) FILING DATE: 04-AUG-1995
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Clark, Paul T.
(B) REGISTRATION NUMBER: 30,162
(C) REFERENCE/DOCKET NUMBER: 07891/002001

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 617/542-5070
(B) TELEFAX: 617/542-8906
(C) TELEX: 200154

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(D) OTHER INFORMATION: Xaa at positons 2, 3, 4, 5,

6, 7, 9, 10, 11, 17, 18, 19, 20, 21, 23, 25, 30, 31, 32, 34, 35, 38, 39, 40, 41, 42, and 45 may be any amino acid. Xaa at positon 8 is Glu or Asp. Xaa at positions 14 & 22 is Val or Ile.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Lys Xaa Cys Met
1               5                   10                  15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Xaa Pro Cys Gly His Xaa Xaa Xaa
            20                  25                  30

Cys Xaa Xaa Cys Ala Xaa Xaa Xaa Xaa Xaa Cys Pro Xaa Cys
        35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(D) OTHER INFORMATION: Xaa at positions 1, 2, 3, 6, 9, 10, 14, 15, 18, 19, 20, 21, 24, 30, 32, 33, 35, 37, 40, 42, 43, 44, 45, 46, 47, 49, 50, 51, 53, 54, 55, 56, 57, 59, 60, 61, 62, 64 and 66 may be any amino acid. Xaa at positions 13, 16 and 17 may be any amino acid or may be absent.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Xaa Xaa Xaa Arg Leu Xaa Thr Phe Xaa Xaa Trp Pro Xaa Xaa Xaa Xaa
1               5                   10                  15

Xaa Xaa Xaa Xaa Xaa Leu Ala Xaa Ala Gly Phe Tyr Tyr Xaa Gly Xaa
            20                  25                  30

Xaa Asp Xaa Val Xaa Cys Phe Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Trp
        35                  40                  45

Xaa Xaa Xaa Asp Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Pro Xaa
    50                  55                  60

Cys Xaa Phe Val
65
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2540 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAAAAGGTGG ACAAGTCCTA TTTTCAAGAG AAGATGACTT TTAACAGTTT TGAAGGATCT 60

AAAACTTGTG TACCTGCAGA CATCAATAAG GAAGAAGAAT TTGTAGAAGA GTTTAATAGA 120

TTAAAAACTT TTGCTAATTT TCCAAGTGGT AGTCCTGTTT CAGCATCAAC ACTGGCACGA 180

GCAGGGTTTC TTTATACTGG TGAAGGAGAT ACCGTGCGGT GCTTTAGTTG TCATGCAGCT 240

GTAGATAGAT GGCAATATGG AGACTCAGCA GTTGGAAGAC ACAGGAAAGT ATCCCCAAAT 300

TGCAGATTTA TCAACGGCTT TTATCTTGAA AATAGTGCCA CGCAGTCTAC AAATTCTGGT 360

ATCCAGAATG GTCAGTACAA AGTTGAAAAC TATCTGGGAA GCAGAGATCA TTTTGCCTTA 420

GACAGGCCAT CTGAGACACA TGCAGACTAT CTTTTGAGAA CTGGGCAGGT TGTAGATATA 480

TCAGACACCA TATACCCGAG GAACCCTGCC ATGTATTGTG AAGAAGCTAG ATTAAAGTCC 540

TTTCAGAACT GGCCAGACTA TGCTCACCTA ACCCCAAGAG AGTTAGCAAG TGCTGGACTC 600

TACTACACAG GTATTGGTGA CCAAGTGCAG TGCTTTTGTT GTGGTGGAAA ACTGAAAAAT 660

TGGGAACCTT GTGATCGTGC CTGGTCAGAA CACAGGCGAC ACTTTCCTAA TTGCTTCTTT 720

GTTTTGGGCC GGAATCTTAA TATTCGAAGT GAATCTGATG CTGTGAGTTC TGATAGGAAT 780

TTCCCAAATT CAACAAATCT TCCAAGAAAT CCATCCATGG CAGATTATGA AGCACGGATC 840

TTTACTTTTG GGACATGGAT ATACTCAGTT AACAAGGAGC AGCTTGCAAG AGCTGGATTT 900

TATGCTTTAG GTGAAGGTGA TAAAGTAAAG TGCTTTCACT GTGGAGGAGG GCTAACTGAT 960

TGGAAGCCCA GTGAAGACCC TTGGGAACAA CATGCTAAAT GGTATCCAGG GTGCAAATAT 1020

CTGTTAGAAC AGAAGGGACA AGAATATATA AACAATATTC ATTTAACTCA TTCACTTGAG 1080

GAGTGTCTGG TAAGAACTAC TGAGAAAACA CCATCACTAA CTAGAAGAAT TGATGATACC 1140

ATCTTCCAAA ATCCTATGGT ACAAGAAGCT ATACGAATGG GGTTCAGTTT CAAGGACATT 1200

AAGAAAATAA TGGAGGAAAA AATTCAGATA TCTGGGAGCA ACTATAAATC ACTTGAGGTT 1260

CTGGTTGCAG ATCTAGTGAA TGCTCAGAAA GACAGTATGC AAGATGAGTC AAGTCAGACT 1320

TCATTACAGA AAGAGATTAG TACTGAAGAG CAGCTAAGGC GCCTGCAAGA GGAGAAGCTT 1380

TGCAAAATCT GTATGGATAG AAATATTGCT ATCGTTTTTG TTCCTTGTGG ACATCTAGTC 1440

ACTTGTAAAC AATGTGCTGA AGCAGTTGAC AAGTGTCCCA TGTGCTACAC AGTCATTACT 1500

TTCAAGCAAA AAATTTTTAT GTCTTAATCT AACTCTATAG TAGGCATGTT ATGTTGTTCT 1560

TATTACCCTG ATTGAATGTG TGATGTGAAC TGACTTTAAG TAATCAGGAT TGAATTCCAT 1620

TAGCATTTGC TACCAAGTAG GAAAAAAAAT GTACATGGCA GTGTTTTAGT TGGCAATATA 1680

ATCTTTGAAT TTCTTGATTT TTCAGGGTAT TAGCTGTATT ATCCATTTTT TTTACTGTTA 1740

TTTAATTGAA ACCATAGACT AAGAATAAGA AGCATCATAC TATAACTGAA CACAATGTGT 1800

ATTCATAGTA TACTGATTTA ATTTCTAAGT GTAAGTGAAT TAATCATCTG GATTTTTTAT 1860

TCTTTTCAGA TAGGCTTAAC AAATGGAGCT TTCTGTATAT AAATGTGGAG ATTAGAGTTA 1920

ATCTCCCCAA TCACATAATT TGTTTTGTGT GAAAAAGGAA TAAATTGTTC CATGCTGGTG 1980

GAAAGATAGA GATTGTTTTT AGAGGTTGGT TGTTGTGTTT TAGGATTCTG TCCATTTTCT 2040

TGTAAAGGGA TAAACACGGA CGTGTGCGAA ATATGTTTGT AAAGTGATTT GCCATTGTTG 2100

AAAGCGTATT TAATGATAGA ATACTATCGA GCCAACATGT ACTGACATGG AAAGATGTCA 2160

GAGATATGTT AAGTGTAAAA TGCAAGTGGC GGGACACTAT GTATAGTCTG AGCCAGATCA 2220

AAGTATGTAT GTTGTTAATA TGCATAGAAC GAGAGATTTG GAAAGATATA CACCAAACTG 2280

TTAAATGTGG TTTCTCTTCG GGGAGGGGGG GATTGGGGGA GGGGCCCCAG AGGGGTTTTA 2340

GAGGGGCCTT TTCACTTTCG ACTTTTTTCA TTTTGTTCTG TTCGGATTTT TTATAAGTAT 2400

GTAGACCCCG AAGGGTTTTA TGGGAACTAA CATCAGTAAC CTAACCCCCG TGACTATCCT 2460

GTGCTCTTCC TAGGGAGCTG TGTTGTTTCC CACCCACCAC CCTTCCCTCT GAACAAATGC 2520

CTGAGTGCTG GGGCACTTTN 2540

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 497 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Thr Phe Asn Ser Phe Glu Gly Ser Lys Thr Cys Val Pro Ala Asp
1               5                  10                  15

Ile Asn Lys Glu Glu Glu Phe Val Glu Glu Phe Asn Arg Leu Lys Thr
            20                  25                  30

Phe Ala Asn Phe Pro Ser Gly Ser Pro Val Ser Ala Ser Thr Leu Ala
        35                  40                  45

Arg Ala Gly Phe Leu Tyr Thr Gly Glu Gly Asp Thr Val Arg Cys Phe
    50                  55                  60

Ser Cys His Ala Ala Val Asp Arg Trp Gln Tyr Gly Asp Ser Ala Val
65                  70                  75                  80

Gly Arg His Arg Lys Val Ser Pro Asn Cys Arg Phe Ile Asn Gly Phe
                85                  90                  95

Tyr Leu Glu Asn Ser Ala Thr Gln Ser Thr Asn Ser Gly Ile Gln Asn
```

```
                100                 105                 110

Gly Gln Tyr Lys Val Glu Asn Tyr Leu Gly Ser Arg Asp His Phe Ala
        115                 120                 125

Leu Asp Arg Pro Ser Glu Thr His Ala Asp Tyr Leu Leu Arg Thr Gly
    130                 135                 140

Gln Val Val Asp Ile Ser Asp Thr Ile Tyr Pro Arg Asn Pro Ala Met
145                 150                 155                 160

Tyr Cys Glu Glu Ala Arg Leu Lys Ser Phe Gln Asn Trp Pro Asp Tyr
                165                 170                 175

Ala His Leu Thr Pro Arg Glu Leu Ala Ser Ala Gly Leu Tyr Tyr Thr
            180                 185                 190

Gly Ile Gly Asp Gln Val Gln Cys Phe Cys Cys Gly Gly Lys Leu Lys
        195                 200                 205

Asn Trp Glu Pro Cys Asp Arg Ala Trp Ser Glu His Arg Arg His Phe
    210                 215                 220

Pro Asn Cys Phe Phe Val Leu Gly Arg Asn Leu Asn Ile Arg Ser Glu
225                 230                 235                 240

Ser Asp Ala Val Ser Ser Asp Arg Asn Phe Pro Asn Ser Thr Asn Leu
                245                 250                 255

Pro Arg Asn Pro Ser Met Ala Asp Tyr Glu Ala Arg Ile Phe Thr Phe
            260                 265                 270

Gly Thr Trp Ile Tyr Ser Val Asn Lys Glu Gln Leu Ala Arg Ala Gly
        275                 280                 285

Phe Tyr Ala Leu Gly Glu Gly Asp Lys Val Lys Cys Phe His Cys Gly
    290                 295                 300

Gly Gly Leu Thr Asp Trp Lys Pro Ser Glu Asp Pro Trp Glu Gln His
305                 310                 315                 320

Ala Lys Trp Tyr Pro Gly Cys Lys Tyr Leu Leu Glu Gln Lys Gly Gln
                325                 330                 335

Glu Tyr Ile Asn Asn Ile His Leu Thr His Ser Leu Glu Glu Cys Leu
            340                 345                 350

Val Arg Thr Thr Glu Lys Thr Pro Ser Leu Thr Arg Arg Ile Asp Asp
        355                 360                 365

Thr Ile Phe Gln Asn Pro Met Val Gln Glu Ala Ile Arg Met Gly Phe
    370                 375                 380

Ser Phe Lys Asp Ile Lys Lys Ile Met Glu Glu Lys Ile Gln Ile Ser
385                 390                 395                 400

Gly Ser Asn Tyr Lys Ser Leu Glu Val Leu Val Ala Asp Leu Val Asn
                405                 410                 415

Ala Gln Lys Asp Ser Met Gln Asp Glu Ser Ser Gln Thr Ser Leu Gln
            420                 425                 430
```

Lys Glu Ile Ser Thr Glu Glu Gln Leu Arg Arg Leu Gln Glu Glu Lys
435 440 445

Leu Cys Lys Ile Cys Met Asp Arg Asn Ile Ala Ile Val Phe Val Pro
450 455 460

Cys Gly His Leu Val Thr Cys Lys Gln Cys Ala Glu Ala Val Asp Lys
465 470 475 480

Cys Pro Met Cys Tyr Thr Val Ile Thr Phe Lys Gln Lys Ile Phe Met
485 490 495

Ser

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2676 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TCCTTGAGAT GTATCAGTAT AGGATTTAGG ATCTCCATGT TGGAACTCTA AATGCATAGA 60

AATGGAAATA ATGGAAATTT TTCATTTTGG CTTTTCAGCC TAGTATTAAA ACTGATAAAA 120

GCAAAGCCAT GCACAAAACT ACCTCCCTAG AGAAAGGCTA GTCCCTTTTC TTCCCCATTC 180

ATTTCATTAT GAACATAGTA GAAAACAGCA TATTCTTATC AAATTTGATG AAAAGCGCCA 240

ACACGTTTGA ACTGAAATAC GACTTGTCAT GTGAACTGTA CCGAATGTCT ACGTATTCCA 300

CTTTTCCTGC TGGGGTTCCT GTCTCAGAAA GGAGTCTTGC TCGTGCTGGT TTCTATTACA 360

CTGGTGTGAA TGACAAGGTC AAATGCTTCT GTTGTGGCCT GATGCTGGAT AACTGGAAAA 420

GAGGAGACAG TCCTACTGAA AAGCATAAAA AGTTGTATCC TAGCTGCAGA TTCGTTCAGA 480

GTCTAAATTC CGTTAACAAC TTGGAAGCTA CCTCTCAGCC TACTTTTCCT TCTTCAGTAA 540

CACATTCCAC ACACTCATTA CTTCCGGGTA CAGAAAACAG TGGATATTTC CGTGGCTCTT 600

ATTCAAACTC TCCATCAAAT CCTGTAAACT CCAGAGCAAA TCAAGAATTT TCTGCCTTGA 660

TGAGAAGTTC CTACCCCTGT CCAATGAATA ACGAAAATGC CAGATTACTT ACTTTTCAGA 720

CATGGCCATT GACTTTTCTG TCGCCAACAG ATCTGGCACG AGCAGGCTTT TACTACATAG 780

GACCTGGAGA CAGAGTGGCT TGCTTTGCCT GTGGTGGAAA ATTGAGCAAT TGGGAACCGA 840

AGGATAATGC TATGTCAGAA CACCTGAGAC ATTTTCCCAA ATGCCCATTT ATAGAAAATC 900

AGCTTCAAGA CACTTCAAGA TACACAGTTT CTAATCTGAG CATGCAGACA CATGCAGCCC 960

GCTTTAAAAC ATTCTTTAAC TGGCCCTCTA GTGTTCTAGT TAATCCTGAG CAGCTTGCAA 1020

GTGCGGGTTT TTATTATGTG GGTAACAGTG ATGATGTCAA ATGCTTTTGC TGTGATGGTG 1080

GACTCAGGTG TTGGGAATCT GGAGATGATC CATGGGTTCA ACATGCCAAG TGGTTTCCAA 1140

GGTGTGAGTA CTTGATAAGA ATTAAAGGAC AGGAGTTCAT CCGTCAAGTT CAAGCCAGTT 1200

ACCCTCATCT ACTTGAACAG CTGCTATCCA CATCAGACAG CCCAGGAGAT GAAAATGCAG 1260

AGTCATCAAT TATCCATTTG GAACCTGGAG AAGACCATTC AGAAGATGCA ATCATGATGA 1320

ATACTCCTGT GATTAATGCT GCCGTGGAAA TGGGCTTTAG TAGAAGCCTG GTAAAACAGA 1380

CAGTTCAGAG AAAAATCCTA GCAACTGGAG AGAATTATAG ACTAGTCAAT GATCTTGTGT 1440

TAGACTTACT CAATGCAGAA GATGAAATAA GGGAAGAGGA GAGAGAAAGA GCAACTGAGG 1500

AAAAAGAATC AAATGATTTA TTATTAATCC GGAAGAATAG AATGGCACTT TTTCAACATT 1560

TGACTTGTGT AATTCCAATC CTGGATAGTC TACTAACTGC CGGAATTATT AATGAACAAG 1620

AACATGATGT TATTAAACAG AAGACACAGA CGTCTTTACA AGCAAGAGAA CTGATTGATA 1680

CGATTTTAGT AAAAGGAAAT ATTGCAGCCA CTGTATTCAG AAACTCTCTG CAAGAAGCTG 1740

AAGCTGTGTT ATATGAGCAT TTATTTGTGC AACAGGACAT AAAATATATT CCCACAGAAG 1800

ATGTTTCAGA TCTACCAGTG GAAGAACAAT TGCGGAGACT ACCAGAAGAA AGAACATGTA 1860

AAGTGTGTAT GGACAAAGAA GTGTCCATAG TGTTTATTCC TTGTGGTCAT CTAGTAGTAT 1920

GCAAAGATTG TGCTCCTTCT TTAAGAAAGT GTCCTATTTG TAGGAGTACA ATCAAGGGTA 1980

CAGTTCGTAC ATTTCTTTCA TGAAGAAGAA CCAAAACATC GTCTAAACTT TAGAATTAAT 2040

TTATTAAATG TATTATAACT TTAACTTTTA TCCTAATTTG GTTTCCTTAA AATTTTTATT 2100

TATTTACAAC TCAAAAAACA TTGTTTTGTG TAACATATTT ATATATGTAT CTAAACCATA 2160

TGAACATATA TTTTTTAGAA ACTAAGAGAA TGATAGGCTT TTGTTCTTAT GAACGAAAAA 2220

GAGGTAGCAC TACAAACACA ATATTCAATC CAAATTTCAG CATTATTGAA ATTGTAAGTG 2280

AAGTAAAACT TAAGATATTT GAGTTAACCT TTAAGAATTT TAAATATTTT GGCATTGTAC 2340

TAATACCGGG AACATGAAGC CAGGTGTGGT GGTATGTACC TGTAGTCCCA GGCTGAGGCA 2400

AGAGAATTAC TTGAGCCCAG GAGTTTGAAT CCATCCTGGG CAGCATACTG AGACCCTGCC 2460

TTTAAAAACN AACAGNACCA AANCCAAACA CCAGGGACAC ATTTCTCTGT CTTTTTTGAT 2520

CAGTGTCCTA TACATCGAAG GTGTGCATAT ATGTTGAATC ACATTTTAGG GACATGGTGT 2580

TTTTATAAAG AATTCTGTGA GNAAAAATTT AATAAAGCAA CCAAATTACT CTTAAAAAAA 2640

AAAAAAAAAA AAAAAACTCG AGGGGCCCGT ACCAAT 2676

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 604 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Asn Ile Val Glu Asn Ser Ile Phe Leu Ser Asn Leu Met Lys Ser
1               5                  10                  15

Ala Asn Thr Phe Glu Leu Lys Tyr Asp Leu Ser Cys Glu Leu Tyr Arg
            20                  25                  30

Met Ser Thr Tyr Ser Thr Phe Pro Ala Gly Val Pro Val Ser Glu Arg
        35                  40                  45

Ser Leu Ala Arg Ala Gly Phe Tyr Tyr Thr Gly Val Asn Asp Lys Val
    50                  55                  60

Lys Cys Phe Cys Cys Gly Leu Met Leu Asp Asn Trp Lys Arg Gly Asp
65                  70                  75                  80

Ser Pro Thr Glu Lys His Lys Lys Leu Tyr Pro Ser Cys Arg Phe Val
                85                  90                  95

Gln Ser Leu Asn Ser Val Asn Asn Leu Glu Ala Thr Ser Gln Pro Thr
            100                 105                 110

Phe Pro Ser Ser Val Thr His Ser Thr His Ser Leu Leu Pro Gly Thr
        115                 120                 125

Glu Asn Ser Gly Tyr Phe Arg Gly Ser Tyr Ser Asn Ser Pro Ser Asn
    130                 135                 140

Pro Val Asn Ser Arg Ala Asn Gln Glu Phe Ser Ala Leu Met Arg Ser
145                 150                 155                 160

Ser Tyr Pro Cys Pro Met Asn Asn Glu Asn Ala Arg Leu Leu Thr Phe
                165                 170                 175

Gln Thr Trp Pro Leu Thr Phe Leu Ser Pro Thr Asp Leu Ala Arg Ala
            180                 185                 190

Gly Phe Tyr Tyr Ile Gly Pro Gly Asp Arg Val Ala Cys Phe Ala Cys
        195                 200                 205

Gly Gly Lys Leu Ser Asn Trp Glu Pro Lys Asp Asn Ala Met Ser Glu
    210                 215                 220

His Leu Arg His Phe Pro Lys Cys Pro Phe Ile Glu Asn Gln Leu Gln
225                 230                 235                 240

Asp Thr Ser Arg Tyr Thr Val Ser Asn Leu Ser Met Gln Thr His Ala
                245                 250                 255

Ala Arg Phe Lys Thr Phe Phe Asn Trp Pro Ser Ser Val Leu Val Asn
            260                 265                 270
```

Pro Glu Gln Leu Ala Ser Ala Gly Phe Tyr Tyr Val Gly Asn Ser Asp
275 280 285

Asp Val Lys Cys Phe Cys Cys Asp Gly Gly Leu Arg Cys Trp Glu Ser
290 295 300

Gly Asp Asp Pro Trp Val Gln His Ala Lys Trp Phe Pro Arg Cys Glu
305 310 315 320

Tyr Leu Ile Arg Ile Lys Gly Gln Glu Phe Ile Arg Gln Val Gln Ala
325 330 335

Ser Tyr Pro His Leu Leu Glu Gln Leu Leu Ser Thr Ser Asp Ser Pro
340 345 350

Gly Asp Glu Asn Ala Glu Ser Ser Ile Ile His Leu Glu Pro Gly Glu
355 360 365

Asp His Ser Glu Asp Ala Ile Met Met Asn Thr Pro Val Ile Asn Ala
370 375 380

Ala Val Glu Met Gly Phe Ser Arg Ser Leu Val Lys Gln Thr Val Gln
385 390 395 400

Arg Lys Ile Leu Ala Thr Gly Glu Asn Tyr Arg Leu Val Asn Asp Leu
405 410 415

Val Leu Asp Leu Leu Asn Ala Glu Asp Glu Ile Arg Glu Glu Glu Arg
420 425 430

Glu Arg Ala Thr Glu Glu Lys Glu Ser Asn Asp Leu Leu Leu Ile Arg
435 440 445

Lys Asn Arg Met Ala Leu Phe Gln His Leu Thr Cys Val Ile Pro Ile
450 455 460

Leu Asp Ser Leu Leu Thr Ala Gly Ile Ile Asn Glu Gln Glu His Asp
465 470 475 480

Val Ile Lys Gln Lys Thr Gln Thr Ser Leu Gln Ala Arg Glu Leu Ile
485 490 495

Asp Thr Ile Leu Val Lys Gly Asn Ile Ala Ala Thr Val Phe Arg Asn
500 505 510

Ser Leu Gln Glu Ala Glu Ala Val Leu Tyr Glu His Leu Phe Val Gln
515 520 525

Gln Asp Ile Lys Tyr Ile Pro Thr Glu Asp Val Ser Asp Leu Pro Val
530 535 540

Glu Glu Gln Leu Arg Arg Leu Pro Glu Glu Arg Thr Cys Lys Val Cys
545 550 555 560

Met Asp Lys Glu Val Ser Ile Val Phe Ile Pro Cys Gly His Leu Val
565 570 575

Val Cys Lys Asp Cys Ala Pro Ser Leu Arg Lys Cys Pro Ile Cys Arg
580 585 590

Ser Thr Ile Lys Gly Thr Val Arg Thr Phe Leu Ser
595 600

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2580 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
TTAGGTTACC TGAAAGAGTT ACTACAACCC CAAAGAGTTG TGTTCTAAGT AGTATCTTGG     60
TAATTCAGAG AGATACTCAT CCTACCTGAA TATAAACTGA GATAAATCCA GTAAAGAAAG    120
TGTAGTAAAT TCTACATAAG AGTCTATCAT TGATTTCTTT TTGTGGTGGA AATCTTAGTT    180
CATGTGAAGA AATTTCATGT GAATGTTTTA GCTATCAAAC AGTACTGTCA CCTACTCATG    240
CACAAAACTG CCTCCCAAAG ACTTTTCCCA GGTCCCTCGT ATCAAAACAT TAAGAGTATA    300
ATGGAAGATA GCACGATCTT GTCAGATTGG ACAAACAGCA ACAAACAAAA AATGAAGTAT    360
GACTTTTCCT GTGAACTCTA CAGAATGTCT ACATATTCAA CTTTCCCCGC CGGGGTGCCT    420
GTCTCAGAAA GGAGTCTTGC TCGTGCTGGT TTTTATTATA CTGGTGTGAA TGACAAGGTC    480
AAATGCTTCT GTTGTGGCCT GATGCTGGAT AACTGGAAAC TAGGAGACAG TCCTATTCAA    540
AAGCATAAAC AGCTATATCC TAGCTGTAGC TTTATTCAGA ATCTGGTTTC AGCTAGTCTG    600
GGATCCACCT CTAAGAATAC GTCTCCAATG AGAAACAGTT TTGCACATTC ATTATCTCCC    660
ACCTTGGAAC ATAGTAGCTT GTTCAGTGGT TCTTACTCCA GCCTTCCTCC AAACCCTCTT    720
AATTCTAGAG CAGTTGAAGA CATCTCTTCA TCGAGGACTA ACCCCTACAG TTATGCAATG    780
AGTACTGAAG AAGCCAGATT TCTTACCTAC CATATGTGGC CATTAACTTT TTTGTCACCA    840
TCAGAATTGG CAAGAGCTGG TTTTTATTAT ATAGGACCTG GAGATAGGGT AGCCTGCTTT    900
GCCTGTGGTG GGAAGCTCAG TAACTGGGAA CCAAAGGATG ATGCTATGTC AGAACACCGG    960
AGGCATTTTC CCAACTGTCC ATTTTTGGAA AATTCTCTAG AAACTCTGAG GTTTAGCATT   1020
TCAAATCTGA GCATGCAGAC ACATGCAGCT CGAATGAGAA CATTTATGTA CTGGCCATCT   1080
AGTGTTCCAG TTCAGCCTGA GCAGCTTGCA AGTGCTGGTT TTTATTATGT GGGTCGCAAT   1140
GATGATGTCA AATGCTTTGG TTGTGATGGT GGCTTGAGGT GTTGGGAATC TGGAGATGAT   1200
CCATGGGTAG AACATGCCAA GTGGTTTCCA AGGTGTGAGT TCTTGATACG AATGAAAGGC   1260
CAAGAGTTTG TTGATGAGAT TCAAGGTAGA TATCCTCATC TTCTTGAACA GCTGTTGTCA   1320
ACTTCAGATA CCACTGGAGA AGAAAATGCT GACCCACCAA TTATTCATTT TGGACCTGGA   1380
GAAAGTTCTT CAGAAGATGC TGTCATGATG AATACACCTG TGGTTAAATC TGCCTTGGAA   1440
```

ATGGGCTTTA ATAGAGACCT GGTGAAACAA ACAGTTCTAA GTAAAATCCT GACAACTGGA 1500

GAGAACTATA AAACAGTTAA TGATATTGTG TCAGCACTTC TTAATGCTGA AGATGAAAAA 1560

AGAGAAGAGG AGAAGGAAAA ACAAGCTGAA GAAATGGCAT CAGATGATTT GTCATTAATT 1620

CGGAAGAACA GAATGGCTCT CTTTCAACAA TTGACATGTG TGCTTCCTAT CCTGGATAAT 1680

CTTTTAAAGG CCAATGTAAT TAATAAACAG GAACATGATA TTATTAAACA AAAAACACAG 1740

ATACCTTTAC AAGCGAGAGA ACTGATTGAT ACCATTTGGG TTAAAGGAAA TGCTGCGGCC 1800

AACATCTTCA AAAACTGTCT AAAAGAAATT GACTCTACAT TGTATAAGAA CTTATTTGTG 1860

GATAAGAATA TGAAGTATAT TCCAACAGAA GATGTTTCAG GTCTGTCACT GGAAGAACAA 1920

TTGAGGAGGT TGCAAGAAGA ACGAACTTGT AAAGTGTGTA TGGACAAAGA AGTTTCTGTT 1980

GTATTTATTC CTTGTGGTCA TCTGGTAGTA TGCCAGGAAT GTGCCCCTTC TCTAAGAAAA 2040

TGCCCTATTT GCAGGGGTAT AATCAAGGGT ACTGTTCGTA CATTTCTCTC TTAAAGAAAA 2100

ATAGTCTATA TTTTAACCTG CATAAAAAGG TCTTTAAAAT ATTGTTGAAC ACTTGAAGCC 2160

ATCTAAAGTA AAAAGGGAAT TATGAGTTTT TCAATTAGTA ACATTCATGT TCTAGTCTGC 2220

TTTGGTACTA ATAATCTTGT TTCTGAAAAG ATGGTATCAT ATATTTAATC TTAATCTGTT 2280

TATTTACAAG GGAAGATTTA TGTTTGGTGA ACTATATTAG TATGTATGTG TACCTAAGGG 2340

AGTAGCGTCN CTGCTTGTTA TGCATCATTT CAGGAGTTAC TGGATTTGTT GTTCTTTCAG 2400

AAAGCTTTGA ANACTAAATT ATAGTGTAGA AAAGAACTGG AAACCAGGAA CTCTGGAGTT 2460

CATCAGAGTT ATGGTGCCGA ATTGTCTTTG GTGCTTTTCA CTTGTGTTTT AAAATAAGGA 2520

TTTTTCTCTT ATTTCTCCCC CTAGTTTGTG AGAAACATCT CAATAAAGTG CTTTAAAAAG 2580

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 618 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met His Lys Thr Ala Ser Gln Arg Leu Phe Pro Gly Pro Ser Tyr Gln
1 5 10 15

Asn Ile Lys Ser Ile Met Glu Asp Ser Thr Ile Leu Ser Asp Trp Thr
20 25 30

Asn Ser Asn Lys Gln Lys Met Lys Tyr Asp Phe Ser Cys Glu Leu Tyr
35 40 45

Arg Met Ser Thr Tyr Ser Thr Phe Pro Ala Gly Val Pro Val Ser Glu
50 55 60

Arg Ser Leu Ala Arg Ala Gly Phe Tyr Tyr Thr Gly Val Asn Asp Lys
65 70 75 80

Val Lys Cys Phe Cys Cys Gly Leu Met Leu Asp Asn Trp Lys Leu Gly
85 90 95

Asp Ser Pro Ile Gln Lys His Lys Gln Leu Tyr Pro Ser Cys Ser Phe
100 105 110

Ile Gln Asn Leu Val Ser Ala Ser Leu Gly Ser Thr Ser Lys Asn Thr
115 120 125

Ser Pro Met Arg Asn Ser Phe Ala His Ser Leu Ser Pro Thr Leu Glu
130 135 140

His Ser Ser Leu Phe Ser Gly Ser Tyr Ser Ser Leu Pro Pro Asn Pro
145 150 155 160

Leu Asn Ser Arg Ala Val Glu Asp Ile Ser Ser Ser Arg Thr Asn Pro
165 170 175

Tyr Ser Tyr Ala Met Ser Thr Glu Glu Ala Arg Phe Leu Thr Tyr His
180 185 190

Met Trp Pro Leu Thr Phe Leu Ser Pro Ser Glu Leu Ala Arg Ala Gly
195 200 205

Phe Tyr Tyr Ile Gly Pro Gly Asp Arg Val Ala Cys Phe Ala Cys Gly
210 215 220

Gly Lys Leu Ser Asn Trp Glu Pro Lys Asp Asp Ala Met Ser Glu His
225 230 235 240

Arg Arg His Phe Pro Asn Cys Pro Phe Leu Glu Asn Ser Leu Glu Thr
245 250 255

Leu Arg Phe Ser Ile Ser Asn Leu Ser Met Gln Thr His Ala Ala Arg
260 265 270

Met Arg Thr Phe Met Tyr Trp Pro Ser Ser Val Pro Val Gln Pro Glu
275 280 285

Gln Leu Ala Ser Ala Gly Phe Tyr Tyr Val Gly Arg Asn Asp Asp Val
290 295 300

Lys Cys Phe Gly Cys Asp Gly Gly Leu Arg Cys Trp Glu Ser Gly Asp
305 310 315 320

Asp Pro Trp Val Glu His Ala Lys Trp Phe Pro Arg Cys Glu Phe Leu
325 330 335

Ile Arg Met Lys Gly Gln Glu Phe Val Asp Glu Ile Gln Gly Arg Tyr
340 345 350

Pro His Leu Leu Glu Gln Leu Leu Ser Thr Ser Asp Thr Thr Gly Glu
355 360 365

Glu Asn Ala Asp Pro Pro Ile Ile His Phe Gly Pro Gly Glu Ser Ser
370 375 380

Ser Glu Asp Ala Val Met Met Asn Thr Pro Val Val Lys Ser Ala Leu
385 390 395 400

Glu Met Gly Phe Asn Arg Asp Leu Val Lys Gln Thr Val Leu Ser Lys
405 410 415

Ile Leu Thr Thr Gly Glu Asn Tyr Lys Thr Val Asn Asp Ile Val Ser
420 425 430

Ala Leu Leu Asn Ala Glu Asp Glu Lys Arg Glu Glu Glu Lys Glu Lys
435 440 445

Gln Ala Glu Glu Met Ala Ser Asp Asp Leu Ser Leu Ile Arg Lys Asn
450 455 460

Arg Met Ala Leu Phe Gln Gln Leu Thr Cys Val Leu Pro Ile Leu Asp
465 470 475 480

Asn Leu Leu Lys Ala Asn Val Ile Asn Lys Gln Glu His Asp Ile Ile
485 490 495

Lys Gln Lys Thr Gln Ile Pro Leu Gln Ala Arg Glu Leu Ile Asp Thr
500 505 510

Ile Trp Val Lys Gly Asn Ala Ala Ala Asn Ile Phe Lys Asn Cys Leu
515 520 525

Lys Glu Ile Asp Ser Thr Leu Tyr Lys Asn Leu Phe Val Asp Lys Asn
530 535 540

Met Lys Tyr Ile Pro Thr Glu Asp Val Ser Gly Leu Ser Leu Glu Glu
545 550 555 560

Gln Leu Arg Arg Leu Gln Glu Glu Arg Thr Cys Lys Val Cys Met Asp
565 570 575

Lys Glu Val Ser Val Val Phe Ile Pro Cys Gly His Leu Val Val Cys
580 585 590

Gln Glu Cys Ala Pro Ser Leu Arg Lys Cys Pro Ile Cys Arg Gly Ile
595 600 605

Ile Lys Gly Thr Val Arg Thr Phe Leu Ser
610 615

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2100 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GACACTCTGC TGGGCGGCGG GCCGCCCTCC TCCGGGACCT CCCCTCGGGA ACCGTCGCCC 60

GCGGCGCTTA GTTAGGACTG GAGTGCTTGG CGCGAAAAGG TGGACAAGTC CTATTTTCCA 120

GAGAAGATGA CTTTTAACAG TTTTGAAGGA ACTAGAACTT TTGTACTTGC AGACACCAAT 180

AAGGATGAAG AATTTGTAGA AGAGTTTAAT AGATTAAAAA CATTTGCTAA CTTCCCAAGT 240

AGTAGTCCTG TTTCAGCATC AACATTGGCG CGAGCTGGGT TTCTTTATAC CGGTGAAGGA 300

GACACCGTGC AATGTTTCAG TTGTCATGCG GCAATAGATA GATGGCAGTA TGGAGACTCA 360

GCTGTTGGAA GACACAGGAG AATATCCCCA AATTGCAGAT TTATCAATGG TTTTTATTTT 420

GAAAATGGTG CTGCACAGTC TACAAATCCT GGTATCCAAA ATGGCCAGTA CAAATCTGAA 480

AACTGTGTGG GAAATAGAAA TCCTTTTGCC CCTGACAGGC CACCTGAGAC TCATGCTGAT 540

TATCTCTTGA GAACTGGACA GGTTGTAGAT ATTTCAGACA CCATATACCC GAGGAACCCT 600

GCCATGTGTA GTGAAGAAGC CAGATTGAAG TCATTTCAGA ACTGGCCGGA CTATGCTCAT 660

TTAACCCCCA GAGAGTTAGC TAGTGCTGGC CTCTACTACA CAGGGGCTGA TGATCAAGTG 720

CAATGCTTTT GTTGTGGGGG AAAACTGAAA AATTGGGAAC CCTGTGATCG TGCCTGGTCA 780

GAACACAGGA GACACTTTCC CAATTGCTTT TTTGTTTTGG GCCGGAACGT TAATGTTCGA 840

AGTGAATCTG GTGTGAGTTC TGATAGGAAT TTCCCAAATT CAACAAACTC TCCAAGAAAT 900

CCAGCCATGG CAGAATATGA AGCACGGATC GTTACTTTTG GAACATGGAT ATACTCAGTT 960

AACAAGGAGC AGCTTGCAAG AGCTGGATTT TATGCTTTAG GTGAAGGCGA TAAAGTGAAG 1020

TGCTTCCACT GTGGAGGAGG GCTCACGGAT TGGAAGCCAA GTGAAGACCC CTGGGACCAG 1080

CATGCTAAGT GCTACCCAGG GTGCAAATAC CTATTGGATG AGAAGGGGCA AGAATATATA 1140

AATAATATTC ATTTAACCCA TCCACTTGAG GAATCTTTGG GAAGAACTGC TGAAAAAACA 1200

CCACCGCTAA CTAAAAAAAT CGATGATACC ATCTTCCAGA ATCCTATGGT GCAAGAAGCT 1260

ATACGAATGG GATTTAGCTT CAAGGACCTT AAGAAAACAA TGGAAGAAAA AATCCAAACA 1320

TCCGGGAGCA GCTATCTATC ACTTGAGGTC CTGATTGCAG ATCTTGTGAG TGCTCAGAAA 1380

GATAATACGG AGGATGAGTC AAGTCAAACT TCATTGCAGA AAGACATTAG TACTGAAGAG 1440

CAGCTAAGGC GCCTACAAGA GGAGAAGCTT TCCAAAATCT GTATGGATAG AAATATTGCT 1500

ATCGTTTTTT TTCCTTGTGG ACATCTGGCC ACTTGTAAAC AGTGTGCAGA AGCAGTTGAC 1560

AAATGTCCCA TGTGCTACAC CGTCATTACG TTCAACCAAA AAATTTTTAT GTCTTAGTGG 1620

GGCACCACAT GTTATGTTCT TCTTGCTCTA ATTGAATGTG TAATGGGAGC GAACTTTAAG 1680

TAATCCTGCA TTTGCATTCC ATTAGCATCC TGCTGTTTCC AAATGGAGAC CAATGCTAAC 1740

AGCACTGTTT CCGTCTAAAC ATTCAATTTC TGGATCTTTC GAGTTATCAG CTGTATCATT 1800

TAGCCAGTGT TTTACTCGAT TGAAACCTTA GACAGAGAAG CATTTTATAG CTTTTCACAT 1860

GTATATTGGT AGTACACTGA CTTGATTTCT ATATGTAAGT GAATTCATCA CCTGCATGTT 1920

TCATGCCTTT TGCATAAGCT TAACAAATGG AGTGTTCTGT ATAAGCATGG AGATGTGATG 1980

GAATCTGCCC AATGACTTTA ATTGGCTTAT TGTAAACACG GAAAGAACTG CCCCACGCTG 2040

CTGGGAGGAT AAAGATTGTT TTAGATGCTC ACTTCTGTGT TTTAGGATTC TGCCCATTTA 2100

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 496 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Thr Phe Asn Ser Phe Glu Gly Thr Arg Thr Phe Val Leu Ala Asp
1 5 10 15

Thr Asn Lys Asp Glu Glu Phe Val Glu Glu Phe Asn Arg Leu Lys Thr
20 25 30

Phe Ala Asn Phe Pro Ser Ser Ser Pro Val Ser Ala Ser Thr Leu Ala
35 40 45

Arg Ala Gly Phe Leu Tyr Thr Gly Glu Gly Asp Thr Val Gln Cys Phe
50 55 60

Ser Cys His Ala Ala Ile Asp Arg Trp Gln Tyr Gly Asp Ser Ala Val
65 70 75 80

Gly Arg His Arg Arg Ile Ser Pro Asn Cys Arg Phe Ile Asn Gly Phe
85 90 95

Tyr Phe Glu Asn Gly Ala Ala Gln Ser Thr Asn Pro Gly Ile Gln Asn
100 105 110

Gly Gln Tyr Lys Ser Glu Asn Cys Val Gly Asn Arg Asn Pro Phe Ala
115 120 125

Pro Asp Arg Pro Pro Glu Thr His Ala Asp Tyr Leu Leu Arg Thr Gly
130 135 140

Gln Val Val Asp Ile Ser Asp Thr Ile Tyr Pro Arg Asn Pro Ala Met
145 150 155 160

Cys Ser Glu Glu Ala Arg Leu Lys Ser Phe Gln Asn Trp Pro Asp Tyr
165 170 175

Ala His Leu Thr Pro Arg Glu Leu Ala Ser Ala Gly Leu Tyr Tyr Thr
180 185 190

Gly Ala Asp Asp Gln Val Gln Cys Phe Cys Cys Gly Gly Lys Leu Lys
195 200 205

Asn Trp Glu Pro Cys Asp Arg Ala Trp Ser Glu His Arg Arg His Phe

```
    210                 215                 220

Pro Asn Cys Phe Phe Val Leu Gly Arg Asn Val Asn Val Arg Ser Glu
225                 230                 235                 240

Ser Gly Val Ser Ser Asp Arg Asn Phe Pro Asn Ser Thr Asn Ser Pro
                245                 250                 255

Arg Asn Pro Ala Met Ala Glu Tyr Glu Ala Arg Ile Val Thr Phe Gly
            260                 265                 270

Thr Trp Ile Tyr Ser Val Asn Lys Glu Gln Leu Ala Arg Ala Gly Phe
        275                 280                 285

Tyr Ala Leu Gly Glu Gly Asp Lys Val Lys Cys Phe His Cys Gly Gly
    290                 295                 300

Gly Leu Thr Asp Trp Lys Pro Ser Glu Asp Pro Trp Asp Gln His Ala
305                 310                 315                 320

Lys Cys Tyr Pro Gly Cys Lys Tyr Leu Leu Asp Glu Lys Gly Gln Glu
                325                 330                 335

Tyr Ile Asn Asn Ile His Leu Thr His Pro Leu Glu Glu Ser Leu Gly
            340                 345                 350

Arg Thr Ala Glu Lys Thr Pro Pro Leu Thr Lys Lys Ile Asp Asp Thr
        355                 360                 365

Ile Phe Gln Asn Pro Met Val Gln Glu Ala Ile Arg Met Gly Phe Ser
    370                 375                 380

Phe Lys Asp Leu Lys Lys Thr Met Glu Glu Lys Ile Gln Thr Ser Gly
385                 390                 395                 400

Ser Ser Tyr Leu Ser Leu Glu Val Leu Ile Ala Asp Leu Val Ser Ala
                405                 410                 415

Gln Lys Asp Asn Thr Glu Asp Glu Ser Ser Gln Thr Ser Leu Gln Lys
            420                 425                 430

Asp Ile Ser Thr Glu Glu Gln Leu Arg Arg Leu Gln Glu Glu Lys Leu
        435                 440                 445

Ser Lys Ile Cys Met Asp Arg Asn Ile Ala Ile Val Phe Phe Pro Cys
    450                 455                 460

Gly His Leu Ala Thr Cys Lys Gln Cys Ala Glu Ala Val Asp Lys Cys
465                 470                 475                 480

Pro Met Cys Tyr Thr Val Ile Thr Phe Asn Gln Lys Ile Phe Met Ser
                485                 490                 495
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Lys Ala Ala Arg Leu Gly Thr Tyr Thr Asn Trp Pro Val Gln Phe Leu
1 5 10 15

Glu Pro Ser Arg Met Ala Ala Ser Gly Phe Tyr Tyr Leu Gly Arg Gly
20 25 30

Asp Glu Val Arg Cys Ala Phe Cys Lys Val Glu Ile Thr Asn Trp Val
35 40 45

Arg Gly Asp Asp Pro Glu Thr Asp His Lys Arg Trp Ala Pro Gln Cys
50 55 60

Pro Phe Val
65

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 275 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Ser Asp Leu Arg Leu Glu Glu Val Arg Leu Asn Thr Phe Glu Lys
1 5 10 15

Trp Pro Val Ser Phe Leu Ser Pro Glu Thr Met Ala Lys Asn Gly Phe
20 25 30

Tyr Tyr Leu Gly Arg Ser Asp Glu Val Arg Cys Ala Phe Cys Lys Val
35 40 45

Glu Ile Met Arg Trp Lys Glu Gly Glu Asp Pro Ala Ala Asp His Lys
50 55 60

Lys Trp Ala Pro Gln Cys Pro Phe Val Lys Gly Ile Asp Val Cys Gly
65 70 75 80

Ser Ile Val Thr Thr Asn Asn Ile Gln Asn Thr Thr Thr His Asp Thr
85 90 95

Ile Ile Gly Pro Ala His Pro Lys Tyr Ala His Glu Ala Ala Arg Val
100 105 110

Lys Ser Phe His Asn Trp Pro Arg Cys Met Lys Gln Arg Pro Glu Gln
115 120 125

Met Ala Asp Ala Gly Phe Phe Tyr Thr Gly Tyr Gly Asp Asn Thr Lys

130 135 140

Cys Phe Tyr Cys Asp Gly Gly Leu Lys Asp Trp Glu Pro Glu Asp Val
145 150 155 160

Pro Trp Glu Gln His Val Arg Trp Phe Asp Arg Cys Ala Tyr Val Gln
165 170 175

Leu Val Lys Gly Arg Asp Tyr Val Gln Lys Val Ile Thr Glu Ala Cys
180 185 190

Val Leu Pro Gly Glu Asn Thr Thr Val Ser Thr Ala Ala Pro Val Ser
195 200 205

Glu Pro Ile Pro Glu Thr Lys Ile Glu Lys Glu Pro Gln Val Glu Asp
210 215 220

Ser Lys Leu Cys Lys Ile Cys Tyr Val Glu Glu Cys Ile Val Cys Phe
225 230 235 240

Val Pro Cys Gly His Val Val Ala Cys Ala Lys Cys Ala Leu Ser Val
245 250 255

Asp Lys Cys Pro Met Cys Arg Lys Ile Val Thr Ser Val Leu Lys Val
260 265 270

Tyr Phe Ser
275

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 498 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Thr Glu Leu Gly Met Glu Leu Glu Ser Val Arg Leu Ala Thr Phe
1 5 10 15

Gly Glu Trp Pro Leu Asn Ala Pro Val Ser Ala Glu Asp Leu Val Ala
20 25 30

Asn Gly Phe Phe Ala Thr Gly Lys Trp Leu Glu Ala Glu Cys His Phe
35 40 45

Cys His Val Arg Ile Asp Arg Trp Glu Tyr Gly Asp Gln Val Ala Glu
50 55 60

Arg His Arg Arg Ser Ser Pro Ile Cys Ser Met Val Leu Ala Pro Asn
65 70 75 80

His Cys Gly Asn Val Pro Arg Ser Gln Glu Ser Asp Asn Glu Gly Asn
85 90 95

Ser Val Val Asp Ser Pro Glu Ser Cys Ser Cys Pro Asp Leu Leu Leu
100 105 110

Glu Ala Asn Arg Leu Val Thr Phe Lys Asp Trp Pro Asn Pro Asn Ile
115 120 125

Thr Pro Gln Ala Leu Ala Lys Ala Gly Phe Tyr Tyr Leu Asn Arg Leu
130 135 140

Asp His Val Lys Cys Val Trp Cys Asn Gly Val Ile Ala Lys Trp Glu
145 150 155 160

Lys Asn Asp Asn Ala Phe Glu Glu His Lys Arg Phe Phe Pro Gln Cys
165 170 175

Pro Arg Val Gln Met Gly Pro Leu Ile Glu Phe Ala Thr Gly Lys Asn
180 185 190

Leu Asp Glu Leu Gly Ile Gln Pro Thr Thr Leu Pro Leu Arg Pro Lys
195 200 205

Tyr Ala Cys Val Asp Ala Arg Leu Arg Thr Phe Thr Asp Trp Pro Ile
210 215 220

Ser Asn Ile Gln Pro Ala Ser Ala Leu Ala Gln Ala Gly Leu Tyr Tyr
225 230 235 240

Gln Lys Ile Gly Asp Gln Val Arg Cys Phe His Cys Asn Ile Gly Leu
245 250 255

Arg Ser Trp Gln Lys Glu Asp Glu Pro Trp Phe Glu His Ala Lys Trp
260 265 270

Ser Pro Lys Cys Gln Phe Val Leu Leu Ala Lys Gly Pro Ala Tyr Val
275 280 285

Ser Glu Val Leu Ala Thr Thr Ala Ala Asn Ala Ser Ser Gln Pro Ala
290 295 300

Thr Ala Pro Ala Pro Thr Leu Gln Ala Asp Val Leu Met Asp Glu Ala
305 310 315 320

Pro Ala Lys Glu Ala Leu Thr Leu Gly Ile Asp Gly Gly Val Val Arg
325 330 335

Asn Ala Ile Gln Arg Lys Leu Leu Ser Ser Gly Cys Ala Phe Ser Thr
340 345 350

Leu Asp Glu Leu Leu His Asp Ile Phe Asp Asp Ala Gly Ala Gly Ala
355 360 365

Ala Leu Glu Val Arg Glu Pro Pro Glu Pro Ser Ala Pro Phe Ile Glu
370 375 380

Pro Cys Gln Ala Thr Thr Ser Lys Ala Ala Ser Val Pro Ile Pro Val
385 390 395 400

Ala Asp Ser Ile Pro Ala Lys Pro Gln Ala Ala Glu Ala Val Ser Asn
405 410 415

Ile Ser Lys Ile Thr Asp Glu Ile Gln Lys Met Ser Val Ser Thr Pro
420 425 430

Asn Gly Asn Leu Ser Leu Glu Glu Glu Asn Arg Gln Leu Lys Asp Ala
435 440 445

Arg Leu Cys Lys Val Cys Leu Asp Glu Glu Val Gly Val Val Phe Leu
450 455 460

Pro Cys Gly His Leu Ala Thr Cys Asn Gln Cys Ala Pro Ser Val Ala
465 470 475 480

Asn Cys Pro Met Cys Arg Ala Asp Ile Lys Gly Phe Val Arg Thr Phe
485 490 495

Leu Ser

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Glu Glu Val Arg Leu Asn Thr Phe Glu Lys Trp Pro Val Ser Phe Leu
1 5 10 15

Ser Pro Glu Thr Met Ala Lys Asn Gly Phe Tyr Tyr Leu Gly Arg Ser
20 25 30

Asp Glu Val Arg Cys Ala Phe Cys Lys Val Glu Ile Met Arg Trp Lys
35 40 45

Glu Gly Glu Asp Pro Ala Ala Asp His Lys Lys Trp Ala Pro Gln Cys
50 55 60

Pro Phe Val
65

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Glu Ala Asn Arg Leu Val Thr Phe Lys Asp Trp Pro Asn Pro Asn Ile
1 5 10 15

Thr Pro Gln Ala Leu Ala Lys Ala Gly Phe Tyr Tyr Leu Asn Arg Leu
20 25 30

Asp His Val Lys Cys Val Trp Cys Asn Gly Val Ile Ala Lys Trp Glu
35 40 45

Lys Asn Asp Asn Ala Phe Glu Glu His Lys Arg Phe Phe Pro Gln Cys
50 55 60

Pro Arg Val
65

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Glu Phe Asn Arg Leu Lys Thr Phe Ala Asn Phe Pro Ser Ser Ser Pro
1 5 10 15

Val Ser Ala Ser Thr Leu Ala Arg Ala Gly Phe Leu Tyr Thr Gly Glu
20 25 30

Gly Asp Thr Val Gln Cys Phe Ser Cys His Ala Ala Ile Asp Arg Trp
35 40 45

Gln Tyr Gly Asp Ser Ala Val Gly Arg His Arg Arg Ile Ser Pro Asn
50 55 60

Cys Arg Phe Ile
65

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Glu Phe Asn Arg Leu Lys Thr Phe Ala Asn Phe Pro Ser Gly Ser Pro
1 5 10 15

Val Ser Ala Ser Thr Leu Ala Arg Ala Gly Phe Leu Tyr Thr Gly Glu
20 25 30

Gly Asp Thr Val Arg Cys Phe Ser Cys His Ala Ala Val Asp Arg Trp
35 40 45

Gln Tyr Gly Asp Ser Ala Val Gly Arg His Arg Lys Val Ser Pro Asn
50 55 60

Cys Arg Phe Ile
65

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Glu Leu Tyr Arg Met Ser Thr Tyr Ser Thr Phe Pro Ala Gly Val Pro
1 5 10 15

Val Ser Glu Arg Ser Leu Ala Arg Ala Gly Phe Tyr Tyr Thr Gly Val
20 25 30

Asn Asp Lys Val Lys Cys Phe Cys Cys Gly Leu Met Leu Asp Asn Trp
35 40 45

Lys Arg Gly Asp Ser Pro Thr Glu Lys His Lys Lys Leu Tyr Pro Ser
50 55 60

Cys Arg Phe Val
65

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Glu Leu Tyr Arg Met Ser Thr Tyr Ser Thr Phe Pro Ala Gly Val Pro
1 5 10 15

Val Ser Glu Arg Ser Leu Ala Arg Ala Gly Phe Tyr Tyr Thr Gly Val
20 25 30

Asn Asp Lys Val Lys Cys Phe Cys Cys Gly Leu Met Leu Asp Asn Trp
35 40 45

Lys Leu Gly Asp Ser Pro Ile Gln Lys His Lys Gln Leu Tyr Pro Ser
50 55 60

Cys Ser Phe Ile
65

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Glu Glu Ala Arg Leu Lys Ser Phe Gln Asn Trp Pro Asp Tyr Ala His
1 5 10 15

Leu Thr Pro Arg Glu Leu Ala Ser Ala Gly Leu Tyr Tyr Thr Gly Ala
20 25 30

Asp Asp Gln Val Gln Cys Phe Cys Cys Gly Gly Lys Leu Lys Asn Trp
35 40 45

Glu Pro Cys Asp Arg Ala Trp Ser Glu His Arg Arg His Phe Pro Asn
50 55 60

Cys Phe Phe Val
65

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Glu Glu Ala Arg Leu Lys Ser Phe Gln Asn Trp Pro Asp Tyr Ala His
1 5 10 15

Leu Thr Pro Arg Glu Leu Ala Ser Ala Gly Leu Tyr Tyr Thr Gly Ile
20 25 30

Gly Asp Gln Val Gln Cys Phe Cys Cys Gly Gly Lys Leu Lys Asn Trp
35 40 45

Glu Pro Cys Asp Arg Ala Trp Ser Glu His Arg Arg His Phe Pro Asn
50 55 60

Cys Phe Phe Val
65

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Glu Asn Ala Arg Leu Leu Thr Phe Gln Thr Trp Pro Leu Thr Phe Leu
1 5 10 15

Ser Pro Thr Asp Leu Ala Arg Ala Gly Phe Tyr Tyr Ile Gly Pro Gly
20 25 30

Asp Arg Val Ala Cys Phe Ala Cys Gly Gly Lys Leu Ser Asn Trp Glu
35 40 45

Pro Lys Asp Asn Ala Met Ser Glu His Leu Arg His Phe Pro Lys Cys
50 55 60

Pro Phe Ile
65

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Glu Glu Ala Arg Phe Leu Thr Tyr His Met Trp Pro Leu Thr Phe Leu
1 5 10 15

Ser Pro Ser Glu Leu Ala Arg Ala Gly Phe Tyr Tyr Ile Gly Pro Gly
20 25 30

Asp Arg Val Ala Cys Phe Ala Cys Gly Gly Lys Leu Ser Asn Trp Glu
35 40 45

Pro Lys Asp Asp Ala Met Ser Glu His Arg Arg His Phe Pro Asn Cys
50 55 60

Pro Phe Leu
65

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Tyr Glu Ala Arg Ile Val Thr Phe Gly Thr Trp Ile Tyr Ser Val Asn
1 5 10 15

Lys Glu Gln Leu Ala Arg Ala Gly Phe Tyr Ala Leu Gly Glu Gly Asp
20 25 30

Lys Val Lys Cys Phe His Cys Gly Gly Gly Leu Thr Asp Trp Lys Pro
35 40 45

Ser Glu Asp Pro Trp Asp Gln His Ala Lys Cys Tyr Pro Gly Cys Lys
50 55 60

Tyr Leu
65

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Tyr Glu Ala Arg Ile Phe Thr Phe Gly Thr Trp Ile Tyr Ser Val Asn
1 5 10 15

Lys Glu Gln Leu Ala Arg Ala Gly Phe Tyr Ala Leu Gly Glu Gly Asp
20 25 30

Lys Val Lys Cys Phe His Cys Gly Gly Gly Leu Thr Asp Trp Lys Pro
35 40 45

Ser Glu Asp Pro Trp Glu Gln His Ala Lys Trp Tyr Pro Gly Cys Lys
50 55 60

Tyr Leu
65

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
His Ala Ala Arg Phe Lys Thr Phe Phe Asn Trp Pro Ser Ser Val Leu
1               5                   10                  15

Val Asn Pro Glu Gln Leu Ala Ser Ala Gly Phe Tyr Tyr Val Gly Asn
            20                  25                  30

Ser Asp Asp Val Lys Cys Phe Cys Cys Asp Gly Gly Leu Arg Cys Trp
        35                  40                  45

Glu Ser Gly Asp Asp Pro Trp Val Gln His Ala Lys Trp Phe Pro Arg
    50                  55                  60

Cys Glu Tyr Leu
65
```

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
His Ala Ala Arg Met Arg Thr Phe Met Tyr Trp Pro Ser Ser Val Pro
1               5                   10                  15

Val Gln Pro Glu Gln Leu Ala Ser Ala Gly Phe Tyr Tyr Val Gly Arg
            20                  25                  30

Asn Asp Asp Val Lys Cys Phe Gly Cys Asp Gly Gly Leu Arg Cys Trp
        35                  40                  45

Glu Ser Gly Asp Asp Pro Trp Val Glu His Ala Lys Trp Phe Pro Arg
    50                  55                  60

Cys Glu Phe Leu
65
```

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Glu Ala Ala Arg Leu Arg Thr Phe Ala Glu Trp Pro Arg Gly Leu Lys
1   5   10   15

Gln Arg Pro Glu Glu Leu Ala Glu Ala Gly Phe Phe Tyr Thr Gly Gln
            20   25   30

Gly Asp Lys Thr Arg Cys Phe Cys Cys Asp Gly Gly Leu Lys Asp Trp
        35   40   45

Glu Pro Asp Asp Ala Pro Trp Gln Gln His Ala Arg Trp Tyr Asp Arg
    50   55   60

Cys Glu Tyr Val
65

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Glu Ala Ala Arg Val Lys Ser Phe His Asn Trp Pro Arg Cys Met Lys
1   5   10   15

Gln Arg Pro Glu Gln Met Ala Asp Ala Gly Phe Phe Tyr Thr Gly Tyr
            20   25   30

Gly Asp Asn Thr Lys Cys Phe Tyr Cys Asp Gly Gly Leu Lys Asp Trp
        35   40   45

Glu Pro Glu Asp Val Pro Trp Glu Gln His Val Arg Trp Phe Asp Arg
    50   55   60

Cys Ala Tyr Val
65

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```
Val Asp Ala Arg Leu Arg Thr Phe Thr Asp Trp Pro Ile Ser Asn Ile
1               5                  10                  15

Gln Pro Ala Ser Ala Leu Ala Gln Ala Gly Leu Tyr Tyr Gln Lys Ile
            20                  25                  30

Gly Asp Gln Val Arg Cys Phe His Cys Asn Ile Gly Leu Arg Ser Trp
        35                  40                  45

Gln Lys Glu Asp Glu Pro Trp Phe Glu His Ala Lys Trp Ser Pro Lys
    50                  55                  60

Cys Gln Phe Val
65
```

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
Glu Ser Val Arg Leu Ala Thr Phe Gly Glu Trp Pro Leu Asn Ala Pro
1               5                  10                  15

Val Ser Ala Glu Asp Leu Val Ala Asn Gly Phe Phe Gly Thr Trp Met
            20                  25                  30

Glu Ala Glu Cys Asp Phe Cys His Val Arg Ile Asp Arg Trp Glu Tyr
        35                  40                  45

Gly Asp Leu Val Ala Glu Arg His Arg Arg Ser Ser Pro Ile Cys Ser
    50                  55                  60

Met Val
65
```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Glu Gln Leu Arg Arg Leu Gln Glu Glu Arg Thr Cys Lys Val Cys Met
1 5 10 15

Asp Lys Glu Val Ser Val Val Phe Ile Pro Cys Gly His Leu Val Val
20 25 30

Cys Gln Glu Cys Ala Pro Ser Leu Arg Lys Cys Pro Ile Cys
35 40 45

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Glu Gln Leu Arg Arg Leu Pro Glu Glu Arg Thr Cys Lys Val Cys Met
1 5 10 15

Asp Lys Glu Val Ser Ile Val Phe Ile Pro Cys Gly His Leu Val Val
20 25 30

Cys Lys Asp Cys Ala Pro Ser Leu Arg Lys Cys Pro Ile Cys
35 40 45

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Glu Gln Leu Arg Arg Leu Gln Glu Glu Lys Leu Ser Lys Ile Cys Met
1 5 10 15

Asp Arg Asn Ile Ala Ile Val Phe Phe Pro Cys Gly His Leu Ala Thr
20 25 30

Cys Lys Gln Cys Ala Glu Ala Val Asp Lys Cys Pro Met Cys

```
35                          40                          45
```

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

```
Glu Gln Leu Arg Arg Leu Gln Glu Glu Lys Leu Cys Lys Ile Cys Met
1               5                   10                  15

Asp Arg Asn Ile Ala Ile Val Phe Val Pro Cys Gly His Leu Val Thr
            20                  25                  30

Cys Lys Gln Cys Ala Glu Ala Val Asp Lys Cys Pro Met Cys
        35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```
Glu Glu Asn Arg Gln Leu Lys Asp Ala Arg Leu Cys Lys Val Cys Leu
1               5                   10                  15

Asp Glu Glu Val Gly Val Val Phe Leu Pro Cys Gly His Leu Ala Thr
            20                  25                  30

Cys Asn Gln Cys Ala Pro Ser Val Ala Asn Cys Pro Met Cys
        35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Glu Lys Glu Pro Gln Val Glu Asp Ser Lys Leu Cys Lys Ile Cys Tyr
1 5 10 15

Val Glu Glu Cys Ile Val Cys Phe Val Pro Cys Gly His Val Val Ala
20 25 30

Cys Ala Lys Cys Ala Leu Ser Val Asp Lys Cys Pro Met Cys
35 40 45

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Ala Val Glu Ala Glu Val Ala Asp Asp Arg Leu Cys Lys Ile Cys Leu
1 5 10 15

Gly Ala Glu Lys Thr Val Cys Phe Val Pro Cys Gly His Val Val Ala
20 25 30

Cys Gly Lys Cys Ala Ala Gly Val Thr Thr Cys Pro Val Cys
35 40 45

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2474 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GAATTCCGGG AGACCTACAC CCCCGGAGAT CAGAGGTCAT TGCTGGCGTT CAGAGCCTAG 60

GAAGTGGGCT GCGGTATCAG CCTAGCAGTA AAACCGACCA GAAGCCATGC ACAAAACTAC 120

ATCCCCAGAG AAAGACTTGT CCCTTCCCCT CCCTGTCATC TCACCATGAA CATGGTTCAA 180

GACAGCGCCT TTCTAGCCAA GCTGATGAAG AGTGCTGACA CCTTTGAGTT GAAGTATGAC 240

TTTTCCTGTG AGCTGTACCG ATTGTCCACG TATTCAGCTT TTCCCAGGGG AGTTCCTGTG 300

TCAGAAAGGA GTCTGGCTCG TGCTGGCTTT TACTACACTG GTGCCAATGA CAAGGTCAAG 360

TGCTTCTGCT GTGGCCTGAT GCTAGACAAC TGGAAACAAG GGGACAGTCC CATGGAGAAG 420

CACAGAAAGT TGTACCCCAG CTGCAACTTT GTACAGACTT TGAATCCAGC CAACAGTCTG 480

GAAGCTAGTC CTCGGCCTTC TCTTCCTTCC ACGGCGATGA GCACCATGCC TTTGAGCTTT 540

GCAAGTTCTG AGAATACTGG CTATTTCAGT GGCTCTTACT CGAGCTTTCC CTCAGACCCT 600

GTGAACTTCC GAGCAAATCA AGATTGTCCT GCTTTGAGCA CAAGTCCCTA CCACTTTGCA 660

ATGAACACAG AGAAGGCCAG ATTACTCACC TATGAAACAT GGCCATTGTC TTTTCTGTCA 720

CCAGCAAAGC TGGCCAAAGC AGGCTTCTAC TACATAGGAC CTGGAGATAG AGTGGCCTGC 780

TTTGCGTGCG ATGGGAAACT GAGCAACTGG GAACGTAAGG ATGATGCTAT GTCAGAGCAC 840

CAGAGGCATT TCCCCAGCTG TCCGTTCTTA AAAGACTTGG GTCAGTCTGC TTCGAGATAC 900

ACTGTCTCTA ACCTGAGCAT GCAGACACAC GCAGCCCGTA TTAGAACATT CTCTAACTGG 960

CCTTCTAGTG CACTAGTTCA TTCCCAGGAA CTTGCAAGTG CGGGCTTTTA TTATACAGGA 1020

CACAGTGATG ATGTCAAGTG TTTATGCTGT GATGGTGGGC TGAGGTGCTG GGAATCTGGA 1080

GATGACCCCT GGGTGGAACA TGCCAAGTGG TTTCCAAGGT GTGAGTACTT GCTCAGAATC 1140

AAAGGCCAAG AATTTGTCAG CCAAGTTCAA GCTGGCTATC CTCATCTACT TGAGCAGCTA 1200

TTATCTACGT CAGACTCCCC AGAAGATGAG AATGCAGACG CAGCAATCGT GCATTTTGGC 1260

CCTGGAGAAA GTTCGGAAGA TGTCGTCATG ATGAGCACGC CTGTGGTTAA AGCAGCCTTG 1320

GAAATGGGCT TCAGTAGGAG CCTGGTGAGA CAGACGGTTC AGTGGCAGAT CCTGGCCACT 1380

GGTGAGAACT ACAGGACCGT CAGTGACCTC GTTATAGGCT TACTCGATGC AGAAGACGAG 1440

ATGAGAGAGG AGCAGATGGA GCAGGCGGCC GAGGAGGAGG AGTCAGATGA TCTAGCACTA 1500

ATCCGGAAGA ACAAAATGGT GCTTTTCCAA CATTTGACGT GTGTGACACC AATGCTGTAT 1560

TGCCTCCTAA GTGCAAGGGC CATCACTGAA CAGGAGTGCA ATGCTGTGAA ACAGAAACCA 1620

CACACCTTAC AAGCAAGCAC ACTGATTGAT ACTGTGTTAG CAAAAGGAAA CACTGCAGCA 1680

ACCTCATTCA GAAACTCCCT TCGGGAAATT GACCCTGCGT TATACAGAGA TATATTTGTG 1740

CAACAGGACA TTAGGAGTCT TCCCACAGAT GACATTGCAG CTCTACCAAT GGAAGAACAG 1800

TTGCGGCCCC TCCCGGAGGA CAGAATGTGT AAAGTGTGTA TGGACCGAGA GGTATCCATC 1860

GTGTTCATTC CCTGTGGCCA TCTGGTCGTG TGCAAAGACT GCGCTCCCTC TCTGAGGAAG 1920

TGTCCCATCT GTAGAGGGAC CATCAAGGGC ACAGTGCGCA CATTTCTCTC CTGAACAAGA 1980

CTAATGGTCC ATGGCTGCAA CTTCAGCCAG GAGGAAGTTC ACTGTCACTC CCAGTTCCAT 2040

TCGGAACTTG AGGCCAGCCT GGATAGCACG AGACACCGCC AAACACACAA ATATAAACAT 2100

GAAAAACTTT TGTCTGAAGT CAAGAATGAA TGAATTACTT ATATAATAAT TTTAATTGGT 2160

TTCCTTAAAA GTGCTATTTG TTCCCAACTC AGAAAATTGT TTTCTGTAAA CATATTTACA 2220

TACTACCTGC ATCTAAAGTA TTCATATATT CATATATTCA GATGTCATGA GAGAGGGTTT 2280

TGTTCTTGTT CCTGAAAAGC TGGTTTATCA TCTGATCAGC ATATACTGCG CAACGGGCAG 2340

GGCTAGAATC CATGAACCAA GCTGCAAAGA TCTCACGCTA AATAAGGCGG AAAGATTTGG 2400

AGAAACGAAA GGAAATTCTT TCCTGTCCAA TGTATACTCT TCAGACTAAT GACCTCTTCC 2460

TATCAAGCCT TCTA 2474

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 602 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

```
Met Asn Met Val Gln Asp Ser Ala Phe Leu Ala Lys Leu Met Lys Ser
1               5                   10                  15

Ala Asp Thr Phe Glu Leu Lys Tyr Asp Phe Ser Cys Glu Leu Tyr Arg
            20                  25                  30

Leu Ser Thr Tyr Ser Ala Phe Pro Arg Gly Val Pro Val Ser Glu Arg
        35                  40                  45

Ser Leu Ala Arg Ala Gly Phe Tyr Tyr Thr Gly Ala Asn Asp Lys Val
    50                  55                  60

Lys Cys Phe Cys Cys Gly Leu Met Leu Asp Asn Trp Lys Gln Gly Asp
65                  70                  75                  80

Ser Pro Met Glu Lys His Arg Lys Leu Tyr Pro Ser Cys Asn Phe Val
                85                  90                  95

Gln Thr Leu Asn Pro Ala Asn Ser Leu Glu Ala Ser Pro Arg Pro Ser
            100                 105                 110

Leu Pro Ser Thr Ala Met Ser Thr Met Pro Leu Ser Phe Ala Ser Ser
        115                 120                 125

Glu Asn Thr Gly Tyr Phe Ser Gly Ser Tyr Ser Ser Phe Pro Ser Asp
    130                 135                 140

Pro Val Asn Phe Arg Ala Asn Gln Asp Cys Pro Ala Leu Ser Thr Ser
145                 150                 155                 160

Pro Tyr His Phe Ala Met Asn Thr Glu Lys Ala Arg Leu Leu Thr Tyr
                165                 170                 175

Glu Thr Trp Pro Leu Ser Phe Leu Ser Pro Ala Lys Leu Ala Lys Ala
            180                 185                 190

Gly Phe Tyr Tyr Ile Gly Pro Gly Asp Arg Val Ala Cys Phe Ala Cys
```

```
        195                 200                 205

Asp Gly Lys Leu Ser Asn Trp Glu Arg Lys Asp Asp Ala Met Ser Glu
    210                 215                 220

His Gln Arg His Phe Pro Ser Cys Pro Phe Leu Lys Asp Leu Gly Gln
225                 230                 235                 240

Ser Ala Ser Arg Tyr Thr Val Ser Asn Leu Ser Met Gln Thr His Ala
                245                 250                 255

Ala Arg Ile Arg Thr Phe Ser Asn Trp Pro Ser Ser Ala Leu Val His
            260                 265                 270

Ser Gln Glu Leu Ala Ser Ala Gly Phe Tyr Tyr Thr Gly His Ser Asp
        275                 280                 285

Asp Val Lys Cys Leu Cys Cys Asp Gly Gly Leu Arg Cys Trp Glu Ser
    290                 295                 300

Gly Asp Asp Pro Trp Val Glu His Ala Lys Trp Phe Pro Arg Cys Glu
305                 310                 315                 320

Tyr Leu Leu Arg Ile Lys Gly Gln Glu Phe Val Ser Gln Val Gln Ala
                325                 330                 335

Gly Tyr Pro His Leu Leu Glu Gln Leu Leu Ser Thr Ser Asp Ser Pro
            340                 345                 350

Glu Asp Glu Asn Ala Asp Ala Ala Ile Val His Phe Gly Pro Gly Glu
        355                 360                 365

Ser Ser Glu Asp Val Val Met Met Ser Thr Pro Val Val Lys Ala Ala
    370                 375                 380

Leu Glu Met Gly Phe Ser Arg Ser Leu Val Arg Gln Thr Val Gln Trp
385                 390                 395                 400

Gln Ile Leu Ala Thr Gly Glu Asn Tyr Arg Thr Val Ser Asp Leu Val
                405                 410                 415

Ile Gly Leu Leu Asp Ala Glu Asp Glu Met Arg Glu Glu Gln Met Glu
            420                 425                 430

Gln Ala Ala Glu Glu Glu Glu Ser Asp Asp Leu Ala Leu Ile Arg Lys
        435                 440                 445

Asn Lys Met Val Leu Phe Gln His Leu Thr Cys Val Thr Pro Met Leu
    450                 455                 460

Tyr Cys Leu Leu Ser Ala Arg Ala Ile Thr Glu Gln Glu Cys Asn Ala
465                 470                 475                 480

Val Lys Gln Lys Pro His Thr Leu Gln Ala Ser Thr Leu Ile Asp Thr
                485                 490                 495

Val Leu Ala Lys Gly Asn Thr Ala Ala Thr Ser Phe Arg Asn Ser Leu
            500                 505                 510

Arg Glu Ile Asp Pro Ala Leu Tyr Arg Asp Ile Phe Val Gln Gln Asp
        515                 520                 525
```

Ile Arg Ser Leu Pro Thr Asp Asp Ile Ala Ala Leu Pro Met Glu Glu
530 535 540

Gln Leu Arg Pro Leu Pro Glu Asp Arg Met Cys Lys Val Cys Met Asp
545 550 555 560

Arg Glu Val Ser Ile Val Phe Ile Pro Cys Gly His Leu Val Val Cys
565 570 575

Lys Asp Cys Ala Pro Ser Leu Arg Lys Cys Pro Ile Cys Arg Gly Thr
580 585 590

Ile Lys Gly Thr Val Arg Thr Phe Leu Ser
595 600

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2416 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CTGTGGTGGA GATCTATTGT CCAAGTGGTG AGAAACTTCA TCTGGAAGTT TAAGCGGTCA 60

GAAATACTAT TACTACTCAT GGACAAAACT GTCTCCCAGA GACTCGCCCA AGGTACCTTA 120

CACCCAAAAA CTTAAACGTA TAATGGAGAA GAGCACAATC TTGTCAAATT GGACAAAGGA 180

GAGCGAAGAA AAAATGAAGT TTGACTTTTC GTGTGAACTC TACCGAATGT CTACATATTC 240

AGCTTTTCCC AGGGGAGTTC CTGTCTCAGA GAGGAGTCTG GCTCGTGCTG GCTTTTATTA 300

TACAGGTGTG AATGACAAAG TCAAGTGCTT CTGCTGTGGC CTGATGTTGG ATAACTGGAA 360

ACAAGGGGAC AGTCCTGTTG AAAAGCACAG ACAGTTCTAT CCCAGCTGCA GCTTTGTACA 420

GACTCTGCTT TCAGCCAGTC TGCAGTCTCC ATCTAAGAAT ATGTCTCCTG TGAAAAGTAG 480

ATTTGCACAT TCGTCACCTC TGGAACGAGG TGGCATTCAC TCCAACCTGT GCTCTAGCCC 540

TCTTAATTCT AGAGCAGTGG AAGACTTCTC ATCAAGGATG GATCCCTGCA GCTATGCCAT 600

GAGTACAGAA GAGGCCAGAT TTCTTACTTA CAGTATGTGG CCTTTAAGTT TTCTGTCACC 660

AGCAGAGCTG GCCAGAGCTG GCTTCTATTA CATAGGGCCT GGAGACAGGG TGGCCTGTTT 720

TGCCTGTGGT GGGAAACTGA GCAACTGGGA ACCAAAGGAT TATGCTATGT CAGAGCACCG 780

CAGACATTTT CCCCACTGTC CATTTCTGGA AAATACTTCA GAAACACAGA GGTTTAGTAT 840

ATCAAATCTA AGTATGCAGA CACACTCTGC TCGATTGAGG ACATTTCTGT ACTGGCCACC 900

TAGTGTTCCT GTTCAGCCCG AGCAGCTTGC AAGTGCTGGA TTCTATTACG TGGATCGCAA 960

TGATGATGTC AAGTGCCTTT GTTGTGATGG TGGCTTGAGA TGTTGGGAAC CTGGAGATGA 1020

CCCCTGGATA GAACACGCCA AATGGTTTCC AAGGTGTGAG TTCTTGATAC GGATGAAGGG 1080

TCAGGAGTTT GTTGATGAGA TTCAAGCTAG ATATCCTCAT CTTCTTGAGC AGCTGTTGTC 1140

CACTTCAGAC ACCCCAGGAG AAGAAAATGC TGACCCTACA GAGACAGTGG TGCATTTTGG 1200

CCCTGGAGAA AGTTCGAAAG ATGTCGTCAT GATGAGCACG CCTGTGGTTA AAGCAGCCTT 1260

GGAAATGGGC TTCAGTAGGA GCCTGGTGAG ACAGACGGTT CAGCGGCAGA TCCTGGCCAC 1320

TGGTGAGAAC TACAGGACCG TCAATGATAT TGTCTCAGTA CTTTTGAATG CTGAAGATGA 1380

GAGAAGAGAA GAGGAGAAGG AAAGACAGAC TGAAGAGATG GCATCAGGTG ACTTATCACT 1440

GATTCGGAAG AATAGAATGG CCCTCTTTCA ACAGTTGACA CATGTCCTTC CTATCCTGGA 1500

TAATCTTCTT GAGGCCAGTG TAATTACAAA ACAGGAACAT GATATTATTA GACAGAAAAC 1560

ACAGATACCC TTACAAGCAA GAGAGCTTAT TGACACCGTT TTAGTCAAGG GAAATGCTGC 1620

AGCCAACATC TTCAAAAACT CTCTGAAGGG AATTGACTCC ACGTTATATG AAAACTTATT 1680

TGTGGAAAAG AATATGAAGT ATATTCCAAC AGAAGACGTT TCAGGCTTGT CATTGGAAGA 1740

GCAGTTGCGG AGATTACAAG AAGAACGAAC TTGCAAAGTG TGTATGGACA GAGAGGTTTC 1800

TATTGTGTTC ATTCCGTGTG GTCATCTAGT AGTCTGCCAG GAATGTGCCC CTTCTCTAAG 1860

GAAGTGCCCC ATCTGCAGGG GGACAATCAA GGGGACTGTG CGCACATTTC TCTCATGAGT 1920

GAAGAATGGT CTGAAAGTAT TGTTGGACAT CAGAAGCTGT CAGAACAAAG AATGAACTAC 1980

TGATTTCAGC TCTTCAGCAG GACATTCTAC TCTCTTTCAA GATTAGTAAT CTTGCTTTAT 2040

GAAGGGTAGC ATTGTATATT TAAGCTTAGT CTGTTGCAAG GGAAGGTCTA TGCTGTTGAG 2100

CTACAGGACT GTGTCTGTTC CAGAGCAGGA GTTGGGATGC TTGCTGTATG TCCTTCAGGA 2160

CTTCTTGGGA TTTGGGAATT TGGGGAAAGC TTTGGAATCC AGTGATGTGG AGCTCAGAAA 2220

TCCTGGAACC AGTGACTCTG GTACTCAGTA GATAGGGTAC CCTGTACTTC TTGGTGCTTT 2280

TCCAGTCTGG GAAATAAGGA GGAATCTGCT GCTGGTAAAA ATTTGCTGGA TGTGAGAAAT 2340

AGATGAAAGT GTTTCGGGTG GGGGCGTGCA TCAGTGTAGT GTGTGCAGGG ATGTATGCAG 2400

GCCAAACACT GTGTAG 2416

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 591 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

```
Met Glu Lys Ser Thr Ile Leu Ser Asn Trp Thr Lys Glu Ser Glu Glu
1               5                  10                  15

Lys Met Lys Phe Asp Phe Ser Cys Glu Leu Tyr Arg Met Ser Thr Tyr
            20                  25                  30

Ser Ala Phe Pro Arg Gly Val Pro Val Ser Glu Arg Ser Leu Ala Arg
        35                  40                  45

Ala Gly Phe Tyr Tyr Thr Gly Val Asn Asp Lys Val Lys Cys Phe Cys
    50                  55                  60

Cys Gly Leu Met Leu Asp Asn Trp Lys Gln Gly Asp Ser Pro Val Glu
65                  70                  75                  80

Lys His Arg Gln Phe Tyr Pro Ser Cys Ser Phe Val Gln Thr Leu Leu
                85                  90                  95

Ser Ala Ser Leu Gln Ser Pro Ser Lys Asn Met Ser Pro Val Lys Ser
            100                 105                 110

Arg Phe Ala His Ser Ser Pro Leu Glu Arg Gly Gly Ile His Ser Asn
        115                 120                 125

Leu Cys Ser Ser Pro Leu Asn Ser Arg Ala Val Glu Asp Phe Ser Ser
    130                 135                 140

Arg Met Asp Pro Cys Ser Tyr Ala Met Ser Thr Glu Glu Ala Arg Phe
145                 150                 155                 160

Leu Thr Tyr Ser Met Trp Pro Leu Ser Phe Leu Ser Pro Ala Glu Leu
                165                 170                 175

Ala Arg Ala Gly Phe Tyr Tyr Ile Gly Pro Gly Asp Arg Val Ala Cys
            180                 185                 190

Phe Ala Cys Gly Gly Lys Leu Ser Asn Trp Glu Pro Lys Asp Tyr Ala
        195                 200                 205

Met Ser Glu His Arg Arg His Phe Pro His Cys Pro Phe Leu Glu Asn
    210                 215                 220

Thr Ser Glu Thr Gln Arg Phe Ser Ile Ser Asn Leu Ser Met Gln Thr
225                 230                 235                 240

His Ser Ala Arg Leu Arg Thr Phe Leu Tyr Trp Pro Pro Ser Val Pro
                245                 250                 255

Val Gln Pro Glu Gln Leu Ala Ser Ala Gly Phe Tyr Tyr Val Asp Arg
            260                 265                 270

Asn Asp Asp Val Lys Cys Leu Cys Cys Asp Gly Gly Leu Arg Cys Trp
        275                 280                 285

Glu Pro Gly Asp Asp Pro Trp Ile Glu His Ala Lys Trp Phe Pro Arg
    290                 295                 300

Cys Glu Phe Leu Ile Arg Met Lys Gly Gln Glu Phe Val Asp Glu Ile
305                 310                 315                 320
```

Gln Ala Arg Tyr Pro His Leu Leu Glu Gln Leu Leu Ser Thr Ser Asp
325 330 335

Thr Pro Gly Glu Glu Asn Ala Asp Pro Thr Glu Thr Val Val His Phe
340 345 350

Gly Pro Gly Glu Ser Ser Lys Asp Val Val Met Met Ser Thr Pro Val
355 360 365

Val Lys Ala Ala Leu Glu Met Gly Phe Ser Arg Ser Leu Val Arg Gln
370 375 380

Thr Val Gln Arg Gln Ile Leu Ala Thr Gly Glu Asn Tyr Arg Thr Val
385 390 395 400

Asn Asp Ile Val Ser Val Leu Leu Asn Ala Glu Asp Glu Arg Arg Glu
405 410 415

Glu Glu Lys Glu Arg Gln Thr Glu Glu Met Ala Ser Gly Asp Leu Ser
420 425 430

Leu Ile Arg Lys Asn Arg Met Ala Leu Phe Gln Gln Leu Thr His Val
435 440 445

Leu Pro Ile Leu Asp Asn Leu Leu Glu Ala Ser Val Ile Thr Lys Gln
450 455 460

Glu His Asp Ile Ile Arg Gln Lys Thr Gln Ile Pro Leu Gln Ala Arg
465 470 475 480

Glu Leu Ile Asp Thr Val Leu Val Lys Gly Asn Ala Ala Ala Asn Ile
485 490 495

Phe Lys Asn Ser Leu Lys Gly Ile Asp Ser Thr Leu Tyr Glu Asn Leu
500 505 510

Phe Val Glu Lys Asn Met Lys Tyr Ile Pro Thr Glu Asp Val Ser Gly
515 520 525

Leu Ser Leu Glu Glu Gln Leu Arg Arg Leu Gln Glu Glu Arg Thr Cys
530 535 540

Lys Val Cys Met Asp Arg Glu Val Ser Ile Val Phe Ile Pro Cys Gly
545 550 555 560

His Leu Val Val Cys Gln Glu Cys Ala Pro Ser Leu Arg Lys Cys Pro
565 570 575

Ile Cys Arg Gly Thr Ile Lys Gly Thr Val Arg Thr Phe Leu Ser
580 585 590